

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2002, 16:44:21 ; Search time 8437.74 Seconds
(without alignments)
6566.332 Million cell updates/sec

Title: US-09-931-157-1

Perfect score: 4105

Sequence: 1 gaattcggcgccctcttg.....taaaagtttacagaaacctt 4105

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: em_estba:*
- 2: em_esthum:*
- 3: em_estin:*
- 4: em_estmu:*
- 5: em_estov:*
- 6: em_estpl:*
- 7: em_estro:*
- 8: em_hic:*
- 9: gb_est1:*
- 10: gb_est2:*
- 11: gb_hic:*
- 12: gb_gss:*
- 13: em_gss_hum:*
- 14: em_gss_inv:*
- 15: em_gss_pln:*
- 16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	797.8	19.4	865	9 AUI18882	AUI18882 AUI18882
2	792.2	19.3	863	9 AUI19546	AUI19546 AUI19546
3	789	19.2	869	9 AUI20519	AUI20519 AUI20519
c 4	725.6	17.7	781	9 AUI936539	AUI936539 wd29b09.x
5	695.8	17.0	764	9 AUI139639	AUI139639 AUI139639
c 6	694.4	16.9	739	9 AUI694278	AUI694278 wd45a01.x
c 7	687.8	16.8	747	9 AUI809396	AUI809396 wf70b01.x
8	683	16.6	742	10 BG116534	BG116534 602317635
9	683	16.6	771	10 BG699589	BG699589 602679314
c 10	673	16.4	677	9 AUI130721	AUI130721 qc16e12.x
c 11	653.4	15.9	724	9 AUI953824	AUI953824 wx69h09.x
12	649.8	15.8	708	9 AUI17099	AUI17099 AUI17099
13	640.8	15.6	1627	11 AK017486	AK017486 Mus muscu
c 14	627	15.3	700	9 AUI913052	AUI913052 tz75f01.x
c 15	621.6	15.1	678	9 AUI804414	AUI804414 tc71a03.x
16	611	14.9	652	10 BF984079	BF984079 602307535
17	603.2	14.7	1043	10 BG114663	BG114663 602315469

18	596.8	14.5	918	10 BF980903	BF980903 602304437
c 19	595.4	14.5	674	9 AUI17092	AUI17092 tq78g05.x
c 20	582.2	14.2	645	10 BE813325	BE813325 RC3-BN003
c 21	569	13.9	609	10 BE328379	BE328379 bs93g06.x
c 22	563.2	13.7	676	10 BG548379	BG548379 602575155
c 23	561.8	13.7	582	9 AUI144480	AUI144480 qb73608.x
c 24	552.6	13.5	572	9 AUI240165	AUI240165 qh39f10.x
c 25	534.8	13.0	595	9 AA234178	AA234178 zt49f02.r
c 26	534.4	13.0	560	9 AUI145859	AUI145859 AUI145859
c 27	524.8	12.8	597	9 AUI145342	AUI145342 AUI145342
c 28	521	12.7	560	9 AA52627	AA52627 zt33f03.r
c 29	520.4	12.7	554	9 AUI252673	AUI252673 qh86c11.x
c 30	519.6	12.7	660	9 AA910453	AA910453 oJ89h09.s
c 31	516.4	12.6	524	9 AUI332625	AUI332625 qg28b02.x
c 32	514.4	12.5	552	9 AUI58110	AUI58110 AUI58110
c 33	513.8	12.5	599	9 AA058926	AA058926 z196f11.r
c 34	513.2	12.5	563	9 AA406195	AA406195 zt24d09.s
c 35	512.4	12.5	664	9 AUI16816	AUI16816 AUI16816
c 36	512.2	12.5	549	9 AUI475501	AUI475501 tJ92h05.x
c 37	512	12.5	542	10 BF084938	BF084938 MRO-FT017
c 38	509.4	12.4	601	10 BF084938	BF084938 MRO-FT017
c 39	508.4	12.4	599	9 AUI130680	AUI130680 xF54f12.x
c 40	507	12.4	552	9 AUI146561	AUI146561 AUI146561
c 41	505.8	12.3	552	9 AW965964	AW965964 EST378037
c 42	502.8	12.2	576	9 AW264427	AW264427 xF03b07.x
c 43	496.2	12.1	579	9 AW993755	AW993755 RC3-BN003
c 44	494.8	12.1	525	9 AA934376	AA934376 op27g11.s
c 45	490.2	11.9	573	9 AUI144020	AUI144020 AUI144020

ALIGNMENTS

RESULT	1	865 bp	mRNA	linear	EST 19-OCT-2000
AUI18882	AUI18882	HEMBA1 Homo sapiens cDNA clone HEMBA1004569 5', mRNA			
LOCUS	AUI18882	sequence.			
DEFINITION	AUI18882				
ACCESSION	AUI18882	GI:10934117			
VERSION	AUI18882.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	1 (bases 1 to 865)				
AUTHORS	Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.				
TITLE	HRI human cDNA project				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3951 Fax: 81-438-52-3952 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute. Location/Qualifiers 1. .865 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="HEMBA1004569" /clone_lib="HEMBA1" /tissue_type="whole embryo, mainly head" /dev_stage="embryo, 10 weeks" /note="Vector: pME18SFL3"				
FEATURES	Source				
	271 a	184 c	158 g	247 t	5 others
BASE COUNT					
ORIGIN					

```
Query Match      19.4%; Score 797.8; DB 9; Length 865;
Best Local Similarity 98.0%; Pred. NO. 3.3e-157;
Matches 847; Conservative 0; Mismatches 12; Indels 5; Gaps 4;

QY 1479 tgaatacgaagaatgacagaacccgatggaattacttagtttcttactgctcattgatt 1538
DB 1 tgnatagcaaatgacagaacccgatggaattacttagtttcttactgctcattgatt 60

QY 1539 acatggtattaaacttggcaaccatgaattcatctatataaaccccatagctctgtatttg 1598
DB 61 acatcggtattaaacttggcaaccatgaattcatctatataaaccccatagctctgtatttg 120

QY 1599 tgagcaagaatttttaaaattgtttccagtcagtcagtcagtcagtcagtcagtcagtc 1658
DB 121 tgagcaagaatttttaaaattgtttccagtcagtcagtcagtcagtcagtcagtcagtc 180

QY 1659 aaagtctgatgcctcgttccccatgaacggaacagcatccagtggaagaaccacgac 1718
DB 161 aaagtctgatgcctcgttccccatgaacggaacagcatccagtggaagaaccacgac 240

QY 1719 aaacaacacacacagaccgagcagccataaggaacagcatgaactgaccaccccttag 1778
DB 241 aaacaacacacacagaccgagcagccataaggaacagcatgaactgaccaccccttag 300

QY 1779 aagcaactcctcgttactcccatgaacggaacagcatccagtggaagaaccacgac 1838
DB 301 aagcaactcctcgttactcccatgaacggaacagcatccagtggaagaaccacgac 360

QY 1839 tcgggaactcctcgttactcccatgaacggaacagcatccagtggaagaaccacgac 1898
DB 361 tcgggaactcctcgttactcccatgaacggaacagcatccagtggaagaaccacgac 420

QY 1899 cttccaaacacgcaa-ggtagactggtttatccaccccaacatctacgaactcgtactt 1957
DB 421 cttccaaacacgcaa-ggtagactggtttatccaccccaacatctacgaactcgtactt 480

QY 1958 ctttaattgatctaatattacatatctcgttcttcttcttcttcttcttcttcttcttct 2017
DB 481 ctttaattgatctaatattacatatctcgttcttcttcttcttcttcttcttcttcttct 540

QY 2018 gctgggggagaatgaagactgttaaatgaacagagagatttactacttttgcata 2077
DB 541 gctgggggagaatgaagactgttaaatgaacagagagatttactacttttgcata 600

QY 2078 aaatagacttttcaagtacatgctgcttcttcttcttcttcttcttcttcttcttct 2137
DB 601 aaatagacttttcaagtacatgctgcttcttcttcttcttcttcttcttcttcttct 660

QY 2138 aactggtcaccatgaaactttagagattaacgcaagaattttctactttttaaagtga- 2196
DB 661 aactggtcaccatgaaactttagagattaacgcaagaattttctactttttaaagtga 720

QY 2197 ttttttctcctcagcaacacaaatggctcagtcagtcagtcagtcagtcagtcagtcag 2256
DB 721 ttttttctcctcagcaacacaaatggctcagtcagtcagtcagtcagtcagtcagtcag 780

QY 2257 tgggtcagctatttttaactgcataatagcctaactg--attattgaacttattta- 2313
DB 781 tgggtcagctatttttaactgcataatagcctaactg--attattgaacttattta- 840

QY 2314 cacatagtttgaaaaaaagac 2337
DB 841 cncntagtttgaaaaaaagac 864

RESULT 2
AUI19546
LOCUS
DEFINITION AUI19546 HEMBA1 Homo sapiens cDNA clone HEMBA1006071 5', mRNA
sequence.
ACCESSION AUI19546
```

```
VERSION AUI19546.1 GI:10934781
EST. human.
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 863)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1539-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix
Research Institute; cDNA library construction: Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
source
1. .863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="HEMBA1006071"
/tissue_type="whole embryo, mainly head"
/dev_stage="embryo, 10 weeks"
/notes="Vector: pME18SFL3"
BASE COUNT 285 a 130 c 130 g 316 t 2 others
ORIGIN
```

```
Query Match      19.3%; Score 792.2; DB 9; Length 863;
Best Local Similarity 97.8%; Pred. NO. 4.9e-156;
Matches 845; Conservative 0; Mismatches 14; Indels 5; Gaps 4;

QY 2045 gaaaccagaagatatttactactcttgcatagaataagagctttcaagtagcatggctag 2104
DB 1 GAAACCAGAAGATATTTACTACTCTTTGCATGAAATAGAGCTTTCAAGTACATGGCTAG 60

QY 2105 cttttatggcagttctcgtgaatttcaatgggaactgggtcacatgaacttttagat 2164
DB 61 CTTTATGGCAGTCTCGTGAATGTTCAATGGGAACGTGTCACCATGAACCTTTAGAT 120

QY 2165 taacgcaagaattcttactcttttttaagtga--tttttgccttcagcaaacacac 2222
DB 121 TAACGACAAGATTTTCTACTCTTTTAAAGTATTTTGTCTTCAGCCAAACACAT 180

QY 2223 atgggtcagtcacttttattgaaatgctatttggtgcccagtttttttaactgcata 2282
DB 181 ATGGGCTCAAGTCACCTTTTATTGAAATGTCATTTGGTGCCAGTATTTTAACTGCA 240

QY 2283 atagcctaactgattatttgaactatttacacatagtttgaaaaaaagacacacac 2342
DB 241 ATAGCCTAACATGATATTGGAACCTTTTACACATAGTTTG-AAAAAAGACAAAAA 299

QY 2343 tagtattcagtgagcaatttagatttagtattttccacgtcactatttttttttaaac 2402
DB 300 TAGTATTCCAGGTGAGCAATTAGATTAGTATTTTTCCACGTCACGTGTTTAAAAAC 359

QY 2403 acaaatctaaagctacacaaataactacagcccttaagcacagctctgtagacacatt 2462
DB 360 ACAAATCTTAAAGCTACAAACAAATACTACAGGCCCTTAAAGCACAGCTCTGATGACACATT 419

QY 2463 tggcagtttaataagatgttactcaagaattttttaagaactgtatttttttttaaat 2522
DB 420 TGGCAGTTTAAATAGATGTTACTCAAGAATTTTAAAGAACGTATTATTATTATTTTAAAT 479

QY 2523 ggtgttttattacaaggagaccttgacatgtttgttatgtttaaatccaaagtaatgctt 2582
```

Db	480	GGTGTATTTACAAAGGACCTTGAACATGTTTGTATGTTAAATTCAAAAGTAATGCTT	539	
Qy	2583	caatcagatagttcttttccacaagttcaatactggttttttcattgtaaaatttttgtagaaa	2642	
Db	540	CAATCAGATAGTCTTTTTCACAAAGTTCAAT-CTGTGTTTTCATGTATGTAATTTTGTATGAAA	598	
Qy	2643	aatacaatgccaagaccacaataatgtaattgtagtgcatttaactctgctgagactttc	2702	
Db	599	AATCAATGTCAAGTACCAAAATGTTAATGTATGTGTGTCATTTAACTCTGCTGAGACTTTC	658	
Qy	2703	agtcactgtatagaagcttaaaacacacactcaagagaaaagatcgaatttttccagat	2762	
Db	659	AGTCACTGTATATAGAAAGTCTTAAACACACCTTAAGAGAAAAGATCGAAATTTTTCAGAT	718	
Qy	2763	gattcggaaattttcattcagggtatttgaatagtgacatatatatgtatatcacatatca	2822	
Db	719	GATTCAGAAATTTTCATTCAGGTATTTGTAAGTAGTACATATATATGATATACATATCA	778	
Qy	2823	cctctattctcttaatttttgtaaagttaactggcagct-aagtccttttttttgatcat	2881	
Db	779	CCTCCTATCTCTTAATTTTCTTAAATGGTAACTGGCAGTAAAGCCTTTTTCACCAT	838	
Qy	2882	tcctttttccatagagaacata	2905	
Db	839	TCCTTTTTCATTTTAGGGAAACTTGA	862	
RESULT 3				
AUI20519 HEMBB1 Homo sapiens cDNA clone HEMBB1000846 5', mRNA				
LOCUS	AUI20519	869 bp	mRNA	linear
DEFINITION	AUI20519	HEMBB1 Homo sapiens cDNA clone HEMBB1000846 5', mRNA	sequence.	EST 19-OCT-2000
ACCESSION	AUI20519			
VERSION	AUI20519.1	GI:10935754		
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	Ora,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.			
TITLE	HRI human cDNA project			
JOURNAL	Unpublished (2000)			
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3951 Fax: 81-438-52-3952 Email: genomics@hri.co.jp			
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.				
FEATURES	Location/Qualifiers			
source	1..869			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
	/clone="HEMBB1000846"			
	/clone_lib="HEMBB1"			
	/tissue_type="whole embryo, mainly body"			
	/dev_stage="embryo, 10 weeks"			
	/note="Vector: pMEL8FL3"			
BASE COUNT	277 a 153 c 148 g 289 t			
ORIGIN	2 others			
Query Match 19.2%; Score 789; DB 9; Length 869;				
Best Local Similarity 98.0%; Pred. No. 2.3e-155;				
Matches 850; Conservative 0; Mismatches 12; Indels 5; Gaps 5;				

QY	1771	acccttagaagcactcctcgttactccataatcctctcgggagaaaaataccacaaggca	1830	
Db	1	ACCCTTAGAAGCACTCCTCGGTACTCCCAATAATCCTCTCGGAGAAAAAATCACAAAGCA	60	
QY	1831	actgtagactccgggaactctctctctgactccctctctcttaattcaactccacacccaag	1890	
Db	61	ACTGTGAGTCCGGGAATCTCTTCTGTGATCCTCTTCTCTTAATTAATCTCCACACCCCAAG	120	
QY	1891	aagaaatgctttccaaaaccgcaa-ggtagactgatttatccaccccaacaactcacaa	1949	
Db	121	AAGAAATGCTTTCCAAAACCGCAAGGTAGACTGGTATTATCCACCACCAACATCTACGAA	180	
QY	1950	tcgtactctcttaattgactctaatttacatatctctgcgtgtgtgatttcagcactaaaaa	2009	
Db	181	TCGTACTCTCTTAATTAATGATCAATTTACATATTTCTCGGTGTGTTATTCAGCACATAAAA	240	
QY	2010	tgtgtagactcggggagagaatgaagactgttaaatgaaccagaaggatatattactact	2069	
Db	241	TGTTGGGAGCTGGGGGAGAAATGAAGACTCTTAATGAACACCAAGAGGATATTACTACTT	300	
QY	2070	ttcatgaaaatagagctttccaagtacatgcttagcttttatggcagttctgggaatgt	2129	
Db	301	TTGCATGAAATAGAGCTTTTCAAGTACATGGCTAGCTTTTATGGCAGTCTCTGGTGAATGT	360	
QY	2130	tcaatgggaactgggtccacctgaaccttttagagatttaacgacaaagatttctactctttt	2189	
Db	361	TCAATGGGAACCTGGTCACCATGAACCTTTAGAGATTACGACAAAGATTTTCTACTTTTT	420	
QY	2190	taagtga-tttttgctcttcagcacaacacacaatatgggctcaggtcactttttttgaa	2248	
Db	421	TAAGTGATTTTTTTGTGCTTCAGCCAAACACAAATATGGCTCAAGTCACTTTTATTGAA	480	
QY	2249	atgcattingtgccagtttttttaactgcataatagccttaacacatgatttattgaaact	2308	
Db	481	ATGTCATTTGGTCCAGTATTTTTTAACATGCATAATAGCCTTAACATGATTAATTTGAAC	540	
QY	2309	atttacacatagtttgaaaaaaagacaaaatagattattcaggtagcgaatttagatta	2368	
Db	541	ATTATACACATAGTTTGAAAAAAGACAAAAAATAGTATTACAGTGAGCAATTAGATTA	600	
QY	2369	gtatttcacgtcacctatttttttttaaacacaaattctaaagctacacaaataac	2428	
Db	601	GTATTTTCCACGTCACTGGTTATTTTTTTAAACACACANATCTAAAGCTACACAAATAC	660	
QY	2429	tacaggcccttaagcacagctctgatgcacatttggcagtttaataagattactcaaa	2488	
Db	661	TACAGGCCCTTAAGCACAGCTCTGATGACACATTTGGCAGTTTAATAGATTCTACTCAA	720	
QY	2489	gaatttttaagaactgtatttttttttaaatggtgttttattacaaggagccttgaa	2548	
Db	721	GAATTTTTTAAGACATGNAATTTTATTTTGTGAATCGTGTGTTTATTTACAAAGGACCTTG-A	779	
QY	2549	catgtttgtatgttaaatcacaagtaatgcttcaatcacatagattctttttcacaaagt	2608	
Db	780	CATGTTTGTGAGGTAATTCAAAAGTAATGCTTTTAATCAGATAGTTC-TTTTCACAAGT	838	
QY	2609	tcaatactgttttccatgtaaaattttg	2635	
Db	839	TCAAT-CTGGTTTTCATGTATATTTGG	864	
RESULT	4			
LOCUS	A1936539/c			
DEFINITION	wd29b09.xl Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone IMAGE:2329529 3', similar to gb:S57498 ENDOTHELIN-1 RECEPTOR PRECURSOR (HUMAN);, mRNA sequence.			
ACCESSION	A1936539			
VERSION	A1936539.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
A1936539 781 bp mRNA linear EST 17-DEC-1999				


```
QY 2442 agcacagtcgtatgacacatttggcagtttaataatagatgttactcaagaattttttaaga 2501
|||||
Db 125 AGCACAGTCGTATGACACATTTGGCAGTTTAAATAGATGTTACTCAAGAATTTTAAAGA 184
QY 2502 actgtatttttttttaaatggtgtttttattacaaaggagacctgaacatgtttttatg 2561
|||||
Db 185 ACTGTATTTTATTTTAAATGCTTTTATTACAAGGACCTTGAACATGTTTGTATG 244
QY 2562 ttaaatcaaaagtaagtcctcaatcagatagttctttttcacaaagtccaactactttt 2621
|||||
Db 245 TTAATTCAAAAGTATGCTTCAATCAGATAGTCTTTTTCACAACTTCAAT-CTGTTT 303
QY 2622 tcatgtaattttgtatgaaaaatcaatgtcaagtaccacaaatgttaatgtatgtcat 2681
|||||
Db 304 TCATGTAATTTTGTATGAAAAATCAATGCAAGTACCAAAATGTTAAATGTATGTGTCAT 363
QY 2682 ttaactctccagagactttcagtcactgtatatagaagtctaaacacacacctaaga 2741
|||||
Db 364 TTAACCTCTGCCGAGACTTTTCAGTGCACCTGTATATAGAAGTCTACACACACCTAAGA 423
QY 2742 aaaaagtcgaatttttcagatgattcggaaatttttcattcaggattttgtaatagtaca 2801
|||||
Db 424 AAAAGATCGAATTTTTCAGATGATTCAGAAATTTTCATTCAGGTATTTGTAATAGTGACA 483
QY 2802 tatatatgtatatacatcacctccctctctctcttcttaatttttggtaaaatgttaactggc 2861
|||||
Db 484 TATATATGTATATACATATCACCTCCTATTTCTTAAATTTTCTTAAAAATGTTAACTGGC 543
QY 2862 agtaagtccttttttgatcattcccttttccatagtaggaacataattttgaaagtggccag 2921
|||||
Db 544 AGTAAGTCCTTTTGTGATTCCTTTCCATATAGGAACATAATTTTGAAGTGGNACAG 603
QY 2922 atgagttatcatgtcagtgaaaaataattaccacaaatgccaccagtaacttaacgat 2981
|||||
Db 604 ATGAGTTTATCATGTGTCAGTGAAAAATAATTACCCACAAATGCCACCAG--AACTTACGAT 661
QY 2982 tcttcaacttctgggggttttcagtagtaacctaaactcccccccccccccaacatctccctccca 3041
|||||
Db 662 TCTCCTACTTCTTGGGGTTTTTCAGTATGAACCTTAACCTCCCCACCCCAACATCTCCCTCCCA 721
QY 3042 cattgtcacatttcaaaaggccc 3065
Db 722 CATTTGCCACCAATTCNAAGGGGNC 745

RESULT 6
AI694278/c
LOCUS
DEFINITION
  wd45a01.xl Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
  IMAGE:2331048 3' similar to gb:S57498 ENDOTHELIN-1 RECEPTOR
  PRECURSOR (HUMAN);, mRNA sequence.
ACCESSION
  AI694278
VERSION
  EST.
KEYWORDS
  EST.
SOURCE
  human.
ORGANISM
  Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
  1 (bases 1 to 739)
  NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
  Tumor Gene Index
  Unpublished (1997)
  Contact: Robert Strausberg, Ph.D.
  Email: cgapbs-r@mail.nih.gov
  This clone is available royalty-free through LLNL; contact the
  IMAGE Consortium (info@image.llnl.gov) for further information.
  Insert Length: 1706 Std Error: 0.00
  Seq primer: -40UP from Gibco
  High quality sequence stop: 458.
FEATURES
  Location/Qualifiers
  1..739
  source
```

```
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2331048"
/clone_lib="Soares_NFL_T_GBC_S1"
/lab_host="DH10B"
/notes="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP-GCB1) were mixed, and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo.
BASE COUNT 252 a 148 c 120 g 219 t
ORIGIN
Query Match 16.9%; Score 694.4; DB 9; Length 739;
Best Local Similarity 98.2%; Pred No. 1.8e-135;
Matches 723; Conservative 0; Mismatches 11; Indels 2; Gaps 2;
QY 3370 cttctgcattcataaaattgtaatcatgttaccattacaaatgggataaaggcagc 3429
|||||
Db 739 CTTCCGCATTCAT-AATCCTGTATTCATGTT-CCCTTCCAAATGGGATATAAGACGACG 682
QY 3430 gtaaaagcagatgagctgtggactagcaataataggtttttgttggttggtttgat 3489
|||||
Db 681 GTCAAGCAGATGAGCTGTGGACTACAAATATAGGGTGTGTTGTTGGTGGTTGAT 622
QY 3490 aaagcagtagttggggctcatattgttctctgtctggagcaaaagcattacacttgaa 3549
|||||
Db 621 AAAGCAGTATGGGGGTGATATGTTTCTGTCTCGACGCAAAAGTCATTACACTTCAA 562
QY 3550 gtattatatgttcttctcctcaatccaatgtgtgatgaaatgccaggtgtctgata 3609
|||||
Db 561 GTATTATATTGTTCTTATCCTCAATGTGGTGATGAAATGGCAGGTTCTCTGATA 502
QY 3610 tttcttcagacttcggcagacagattgtgtataataaattaggaagaataatttttgg 3669
|||||
Db 501 TTTCTTTTCAGACTTCGCCAGACAGATTGCTGATAATAAATAGGTAAAGTAATTTCTTG 442
QY 3670 gccatttttaggacaggtaaaataacatcaggttccagttgcttgaattgcaagctaa 3729
|||||
Db 441 GCCATATTTTAGCACAGGTAAATAACATCAGGTTCCAGTTTGTGTAATTCGAAGCTAA 392
QY 3730 gaagtactgcccccttttggtagcagtcacaaatctatttccactggcgcacatcatg 3789
|||||
Db 381 GAAGTACTCCCCCTTTTGTGTGTAGCAGTCAAACTCTATTATTCACCTGGCGCATCATG 332
QY 3790 cagtgatataatgcctataataataagccataggttccacacacattttgtttagacaattg 3849
|||||
Db 321 CAGTGATATATGCTTATAATAAGCCATPAGGTTTCACACATTTTGTGTTAGACAAATGTC 262
QY 3850 ttttttcaagatgctttgtttcttttcataatgaaaaaaatgcattttataaaatcagaaa 3909
|||||
Db 261 TTTTTCAGATGCTTTGTTCTTTCATATGAAAAAATGCATTTTATAAATTCAGAAA 202
QY 3910 gtcataagttctgaaggcgtcaacgtgcattttatttggactggtaagtaactgttg 3969
|||||
Db 201 GTCATAGATTTCTGAAGGCGTCAACGTCGATTTATTTATGGAGCTGTAAGTAACGTGG 142
QY 3970 tttactagcaggaataatttccaaatttctacccttactacatcttttcaacaagaacttt 4029
|||||
Db 141 TTTACTAGCAGGACTATTTCCAAATTTCTACCTTTTACTACATCTTTTCAACAAGTAACACTT 82
QY 4030 gtagaatgagcagaagccaaaggccctgagttggcagtgggcccatagtgtaaaataaa 4089
|||||
Db 81 GTAGAAATGAGCCAGAACCCAGGGCCCTGAGTTGGCAGTGGGCCATAGTGTAATAATAA 22
```

```

QY 4090 agttacagaaacctt 4105
|||||
Db 21 AGTTTACAGAAACCTT 6

RESULT 7
LOCUS AI809396
DEFINITION wf70b01.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
IMAGE:2360905 3' similar to gb:S57498 ENDOTHELIN-1 RECEPTOR
PRECURSORS (HUMAN);, mRNA sequence.
ACCESSION AI809396
VERSION AI809396
KEYWORDS AI809396.1 GI:5395962
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 747)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1717 Std Error: 0.00
Seq primer: -400P from Gibco
High quality sequence stop: 466.
Location/Qualifiers
1..747
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2360905"
/lab_host="Soares_NFL_T_GBC_S1"
/Note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site:1: Not I; Site_2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19W, testis NHT, and B-cell
NCI-CGAP GCBI) were mixed, and ss circles were made in
vitro. Following RAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo. "
BASE COUNT 250 a 154 c 118 g 221 t 4 others
ORIGIN

Query Match 16.8%; Score 687.8; DB 9; Length 747;
Best Local Similarity 98.1%; Pred. No. 4.4e-134;
Matches 725; Conservative 0; Mismatches 11; Indels 3; Gaps 3;

QY 3368 acctctgcatcataaaattgttaactgttaccattacaataaggatataagggca 3427
|||||
Db 742 ACCCTTGTGATTTCAT-AACTGGTAAATCATG-TACCATACNNAATGGGATATAAGAGGCA 685

QY 3428 gcgtgaaagcagatgagctgtgactagcaataatagggtttgtttgttg-ttggttt 3486
|||||
Db 684 GCGTGAAGCAGATGAGCTGTGGACTAGCAATATAGGGTTTGTGGTGGGTTGGTNT 625

QY 3487 gataagcagattgtggggtcatattgttctgtgctggagcaaaagtcattacacttt 3546
|||||
Db 624 GATAAGCAGTATTGGGGCTCATATTGTTTCTGTCTGTGAGCAAAAGTCATTACACTTT 565

QY 3547 gaagtattattgttcttctatcctcaattcaatgtgtggtgatgaaattgccaggttctgt 3606
|||||
Db 564 GAAGTATTATTATTGTTCTTCTCAATTCATGTGGTATGAAATGGCAGGTGCTGTCG 505

us-09-931-l57-l-1.rst
Page 6

```


JOURNAL
COMMENT

Unpublished (1998)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Chris Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D. cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: Christa Prange, The I.M.A.G.E. Consortium DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www-bio.llnl.gov/bbrp/image/image.html
Seq primer: -400P from Gibco
High quality sequence stop: 420.

FEATURES

source

1. .724
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2548961"
/clone_lib="NCI_CGAP_Brn53"
/tissue_type="three pooled meningiomas"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Library constructed by Life Technologies."
242 a 150 c 110 g 216 t 6 others

BASE COUNT
ORIGIN

Query Match 15.9%; Score 653.4; DB 9; Length 724;
Best Local Similarity 96.4%; Pred. No. 7.5e-127;
Matches 697; Conservative 0; Mismatches 22; Indels 4; Gaps 3;

QY 3384 aatcttgtaacatgctaccattacaaatgggataaagcagcgtgaaagcagatga 3443
|||||
DB 724 AATCTGTAATCATGTTAC - ATCCAAATGGGATTACAGCAGCGTGAAGCAGCATCA 667
|||||

QY 3444 gctgtggaactagcaatatagggttttggttggttggttggttggataaagcagattgg 3503
|||||
DB 666 GCTGTGNACTA - CAAATAGGGTTTGTNNTTGGTTGGTTGTGATGAACAGGTATGGG 608
|||||

QY 3504 gtcataattgtctcgtctgagcaaaagtcattacactttg-aagtattatattgt 3562
|||||
DB 607 GGTGATATTGTTCCGTGCTCGAGCAAAAGTCATTACACTTTGAAGATATANTATTGTT 548
|||||

QY 3563 ctatctcctaattcaatgtgtgtagaataatgccaggttgcgtgatatctttcttcagact 3622
|||||
DB 547 CTATCTCAATTCATGTTGATGAAATTCAGGTTGCTGATATTCTTTTCAGACT 488
|||||

QY 3623 tcgcagacagattgctgataataaattaggtaagataattgttgggccataattttagg 3682
|||||
DB 487 TCNCACAGACAGATTGCTGATAATAATTAGTAAAGATAATTTGTTGGGCCATATTTTAGG 428
|||||

QY 3683 acaggttaaaataacatccaggttccagttgcttgaattgcaaggctaaagactgacct 3742
|||||
DB 427 ACAGGTAAATTAACATCAGGTTCCAGTTCCTGTAATTGCAAGGCTAAGAGTACTGCCCT 368
|||||

QY 3743 ttgtgtgttagcagtcacaaatctatttccactggcgcatacatgcagtgatatatgc 3802
|||||
DB 367 TTTGTGTGTAGCAGTCAAACTATTATTCCACTGGCGCATCATATGCAAGTATATGC 308
|||||

QY 3803 ctataataagccataggttcacaccatttggtttagacaattgtcttttttccaagat 3862
|||||
DB 307 CTATAATATAAGCCATAGGTTTCACACCAATTTGTTTAGACAAATGCTCTTTTTCAGAT 248
|||||

QY 3863 gcttgttctcttcataaataaataatgcattttataaattcagaagcatagattctc 3922
|||||
DB 247 GCTTTGTTCTTCATATGAAAAAATGCATTTTATAAATTCAGAAAGTCATAGATTCT 188
|||||

QY 3923 gaaggcgtcaacgctgattttatttattgagctggtaagtaactgtggttttacttagcagga 3982
|||||
DB 187 GAAGGGCTCAACGTCGATTTTATTATGGACTGGTAAGTAAGTACTGTGTTTACTAGCAGGA 128
|||||

QY 3983 atatttccaaatttctacctttactacatcttttcaacaagtaactttgtagaatagacc 4042
|||||

Db 127 ATATTCCAAATTTCTACTTTTACTATCATCTTTTCAACAAGTAACCTTTGTAGAATGAGCC 68
|||||

QY 4043 agaagccaagggcctgagttggcagtgcccaataaagttaaaataaaagttcacagaac 4102
|||||

Db 67 AGAAGCCAAGGCCCTCAGTTGGCAGTGGCCCATAGCTGTAAATAATAAAGCTTTACAGAAC 8
|||||

QY 4103 ctt 4105
|||||

Db 7 CTT 5
|||||

RESULT 12
AUI17099
LOCUS AUI17099 HEMBA1 Homo sapiens cDNA clone HEMBA1000665 5', mRNA
DEFINITION sequence.
ACCESSION AUI17099
VERSION AUI17099.1 GI:10932041
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 708)
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished (2000)
COMMENT Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3951
Fax: 81-438-52-3952
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.
FEATURES
source
1. .708
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="HEMBA1000665"
/clone_lib="HEMBA1"
/tissue_type="whole embryo, mainly head"
/dev_stage="embryo, 10 weeks"
/note="Vector: pME18SFL3"
BASE COUNT 218 a 126 c 109 g 250 t 5 others
ORIGIN
Query Match 15.8%; Score 649.8; DB 9; Length 708;
Best Local Similarity 97.1%; Pred. No. 4.3e-126;
Matches 691; Conservative 0; Mismatches 17; Indels 4; Gaps 3;

QY 2436 ccttaaaagcacagtcgatgcacacatttggcagtttaataagattactcaagaatttt 2495
|||||

Db 1 CCTTAAAGCACAGCTGATGACACATTTGGCAGTTTAATAGATGTTTACTCAAGAATTTT 60
|||||

QY 2496 ttaagaactgtatt 2555
|||||

Db 61 TTAAGAACGTATTTTATTTTATATGTTGTTTTTATCAAGGGACCTTGAACATGTTT 120
|||||

QY 2556 tgatgttaaatcaaaagtaaatgcctcaatcagatagttctttttcacaaagttcaatc 2615
|||||

Db 121 TGTATGTTAAATTCAAAAGTAATGCTTCAATCAGATAGTCTCTTTTTCACAAGTCAAT-C 179
|||||

QY 2616 tgtttttcatgaaaatttttggatgaataatcaatgcagttaccacaaaattgtaagtatg 2675
|||||

Db 180 TGTTTTTCATGTAATTTTGTATGAAAAATCAATGTCAAGTACCAAAATGTTTAATGTATG 239
|||||

Db 2 AGAGTAGAGGAGGAGGCTGTAGGAGCCTGTGGAGTCTAAGGAAGATC--CGGGAGGCGGTG 59

Qy 88 ttctcccgagttcttcttcgtgagccctcgcgcggtacatcccg-ctg 146

Db 60 TTCTCCGGAGTTGCTTTCTTGGGAGCCT-----CGCGGACACCCATCCCTTTCTA 114

Qy 147 gctgacgatgtgagagcggtgagagagcttccatccaccccggtcgctgcgg 206

Db 115 GTCTGGCAACTGTGTCTAGGAGGTGGGAGGCTCTCTGTATCCACCGACCATCGCTGG 174

Qy 207 gattggggtccagcgacacctcccccgggagagcagtgcccgagaggttttctgagc 266

Db 175 AGCTTGCAAGGCTGACCAAGATCTCC-----CCTAGAGAGCTGGCTGTC 220

Qy 267 cgggaaagtgtgagccgagccgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 326

Db 221 CGGGGAAGTTTCCCG-----AGCTGAGACTGTGCTGCAAGCCCTGCTACCGGCCACCCCTG 276

Qy 327 cagccacccacccctcgttctcgcggttctcgtgcccagcgcgcgccgcgcgcgcgcgc 386

Db 277 CGCGCCACCCCTGCTTCTCCAGCTCAGGCTCCGCTGGCGCTGGCCG-----GGACCTGG 330

Qy 387 agctgtcgcgcagccgagctccacggtgaaaaaaagtgaagtgtaaaagcagcac 446

Db 331 AGCTGTCTGCTTCCGAGGAGCTCTAAGGTGAAAAAAGA---AAGCGGTGAGAGCAACT 387

Qy 447 aagtcaataagagatatcttcccaaatgtcccaagatggaacccctttgcctcagg 506

Db 388 AAG-----AAGACTTAAATCCAGGTTAAGATGAGTATCTTTTGGCTTGGCG 434

Qy 507 cctcttttgcgtggcaactgttgatgtgtaactcagtgataactctgagagatacaga 566

Db 435 CATACTTTGGCTGACCATGTGGGAGCGCTAATGGCTGACAATCCGAGAGATACACGG 494

Qy 567 caactcaagcaatcgtgatgattccaccacttttcggtggcacagagctcagctcc 626

Db 495 CTATCTAAGCAGCAGCATGGAAGACTTCACCCCTTTTCCGGGACGAGATCAACTTTC 554

Qy 627 tggttaccactcacaacccactaatgttggtctaccagcaatggctcactgacacact 686

Db 555 TGG-GCCACCCATCGACCCCTAATTTGGGCGGTGATAGCAATGGCTCAATGCAAGCT 613

Qy 687 attgcccacagacagataaattacttcagctttccaaatacattaaactgtgatatct 746

Db 614 ATTG-CCACAGCAGACTAAATTCAGAGAGCTTTCAGATATATTAACTGTGATATCT 672

Qy 747 gtactatttctcgtgggaatggtggggaatgcaactctgctcaggtatcattaccaga 806

Db 673 GCACCACTTTTCATCGTGGAGTGTGGGGAACGCAACTCTACTAGCAATCATTTATCAA 732

Qy 807 acaatgtatgaggaatggcccaacgcgctgatagccagcttctgccccttgg-agacctt 865

Db 733 ACAAGTGTATGAGG-ACGCGCCCAATGGCTCATAGCCAGCTGGGCTCTTGGAGACCTT 791

Qy 866 attcatgtggtcattgatctccctatacaattatttaagctgctggtggtggtggtcct 925

Db 792 ATCTACGCTGTGATGACCTCCCATCAACCGTGTTA--GCTCTGTGAGGACGCT-GCCT 848

Qy 926 ttgatcacatgacttggcggtattcttttgcaagctgttcccccttttgcagaagtc 985

Db 849 TTCACCAACATGATTTTGGAGTGTNCTCTGCAAGCTGTCCCGCTTCTCTGCAAGATCC 908

Qy 986 tcggtggggatcacgctcctcacaactcgtcctttagtggtagacagacagcagtt 1045

Db 909 TCCGTGGGCATCACCGCTTCAACCTCTGTCTCTCAGTGTGGACAGGTACAGAGCAGTG 968

Qy 1046 gcctcctggagctgtgttcaagggaattgggatttctcttggtaactgacctgaaattgct 1105

Db 969 GCTTCTTGAGCCCGAGTTCAAGGAATCGAGATCCCTTGTATTTACCGCATTTGAATCTGC 1028

Qy 1106 tccatctggatccctcgtcttctctcgtgcccattcctgaaagcagattggtctcgtcatgta 1165

Db 1029 TCCATCTGGATCTTTCTCTCATCTTGGCCATTCGGGAAGCAATCGGCTTCTCATGTGTA 1088

Qy 1166 ccctttgaataggggtgaacagcataaaaacctgtatgtctcaatgcacatacaaatc 1225

Db 1089 CCCTTCGAATACAAGGGCAGCTGCTAGGACCTGCTCAAGCCACCTGCTCAAGTTC 1148

Qy 1226 atgaggttctaccagatgtaaaagactggtggtccttcggttcttatttctgtatgcc 1285

Db 1149 ATGGAGTATTTACCAAGATGTGAAGACTGGTGGCTTTGGGTTCTACTTCTGCAATGCC 1208

Qy 1286 ttggtgtgcactgcgatcttctcacacctcatgactgtgagatgtgaaacagaagaaat 1345

Db 1209 TTGTTGTGCACAGCAATCTTCTACACCCTCATGACCTGTGAGATGCTCAACAGGAGGAAC 1268

Qy 1346 ggcagcttgagaattgcccctcagtgaaacatctttaagcag 1384

Db 1269 GGCAGCTTGGGATCGCCCTTAGTGAGCAGCTCAAAACAG 1307

RESULT 14

A1913052/c

LOCUS

DEFINITION

700 bp mRNA .linear EST 16-DEC-1999

tz75f01.x1 NCI_CGAP_Pan1 Homo sapiens cDNA clone IMAGE:2294425 3'

similar to gb:S57498 ENDOTHELIN-1 RECEPTOR PRECURSOR (HUMAN);, mRNA

sequence.

ACCESSION

A1913052

VERSION

A1913052.1

GI:5632907

KEYWORDS

EST.

SOURCE

human.

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS

NCI-CCGAP (bases 1 to 700)

TITLE

National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL

Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Life Technologies catalog #: 11548-013
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CCGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: www.bio.llnl.gov/bbrp/image/image.html
Insert Length: 3088 Std Error: 0.00
Seq primer: -400p from Gibco
High quality sequence stop: 412.

FEATURES

Location/Qualifiers

1..700

Source

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2294425"
/clone_lib="NCI_CGAP_Pan1"
/tissue_type="adenocarcinoma"
/lab_host="DH10B"
/note="Organ: pancreas; Vector: pCMV-SPORT6; Site1: SalI; Site2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.72 kb. Life Technologies catalog #: 11548-013"

BASE COUNT

232 a 144 c 106 g 214 t 4 others

ORIGIN

Query Match 15.3%; Score 627; DB 9; Length 700;
Best Local Similarity 96.1%; Pred. No. 2.7e-121;
Matches 662; Conservative 0; Mismatches 24; Indels 3; Gaps 2;

Qy 3418 ataagagcagcgtgaagcagcagatgagctggtgactagaataataggtttgttggtt 3477

Db 700 ATATAGACAGNCGTGAAGCAGCAGATGAGCTGGACTACAAAATAGGGTTTGTGGGTG 641

Qy 3478 ggttgggttgataaagcagatttgggtcattattgttctcgtcgtgagcaaaagca 3537

Db 640 GGTGGTTTGAT--AAGCAGTATTGGGGGTGATATTGTTCTCTGCTCGGCAAAAGTCA 583

QY	3538	ttacactttgaa -gtatttatattgtcttctcaattcaatgctgggtgagaaattgcc	3596
Db	582	TNACACTTTGAAGGTATANTATTGTTCTTATCNTCAATTCATGTGGTGATGAATATGCC	523
QY	3597	aggtgtctgatatctcttcagacttcggcagacagattgctgataataaattaggttaa	3656
Db	522	AGGTGTGCTGATATTTCTTTTCAGACTTCGCCGAGACAGATTGCTGATATAAATAGGTAA	463
QY	3657	gataatttgttggcccatatttttaggcagagtaaaataacatcaggttccagttgcttga	3716
Db	462	GNTAATTTGTGTGGCCCATATTTTAGGCAGAGTAAATTAACATCAGGTGCCAGTTCCTGA	403
QY	3717	attcaaggctcaagaagtactgcacctttgtgttagcagtcacaaatctattatccact	3776
Db	402	ATTGCAAGGCTAACAAGTACTGCCTTTTGTGTGTAGCAGTCAAATCTATTATCCACT	343
QY	3777	ggcgcatcatatgcagtgatatatgcctataataataagccataggttcacaccatttgt	3836
Db	342	GGCGCATCATATGCAGTGATATATGCGCTATATAATATAAGCCATAGGTTTCACACCATTTTGT	283
QY	3837	ttagacaattgtcttttttccaagatgctttgtttcttccatgtaaaaaaatgcatttt	3896
Db	282	TTAGACAATTTGCTCTTTTTCACAGATGCTTTGTTCTTTCATATGAAAAAATGCATTTT	223
QY	3897	ataaattcagaagtcatagatattctgaagcgtcaacgtgcattttatttatggactgg	3956
Db	222	ATAAATTCAGAAAGTCATAGATTTCGTAAGCGTCACAGTGCATTTTATTATGACTGG	163
QY	3957	taagtaactgtgttactagcaggaatatttccaatcttccacttactacatcttttc	4016
Db	162	TAGTAACACTGTGGTTTACTAGCAGGAATATTTCCAAATTTCTACCTTTTACATACATCTTTTC	103
QY	4017	aacaagttaactttgtagaatatgagccagaagcccaagccctgagttggcagtgcccata	4076
Db	102	AACAAGTAACTTTGTAGAATAGCCAGAGCCCAAGGCCCTTGAGTTGGCAGTGCCCATTA	43
QY	4077	agtgtaaaataaaagtttacagaaacctt	4105
Db	42	AGTGTAATAATAAAAGTTTACAGAACCCCTT	14

RESULT 15
AI804414/c
LOCUS
DEFINITION
AI804414 678 bp mRNA linear EST 13-DEC-1999
tc7la03.x1 Soares_NhmPpu.S1 Homo sapiens cDNA clone IMAGE:2070028
3' similar to gb:S57498 ENDOTHELIN-1 RECEPTOR PRECURSOR (HUMAN);,
mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 678)
NCI-CCGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 755 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 464.
Location/Qualifiers
1..678
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2070028"
/clone_lib="Soares_NhmPpu.S1"

Search completed: September 30, 2002, 20:31:22

Job time: 13621 sec

Sequence	Strd	Orig	Zscore	EScore	Len	Documentation
sp_mammal:Q95155	+	2156.00	3759.62	1.4e-201	427	! Q95155 ovis aries (sheep). end
sp_rodent:Q91VV2	+	2126.00	3706.92	9.2e-199	427	! Q91VV2 mus musculus (mouse). s
sp_mammal:Q9GL61	+	2046.00	3567.28	6.0e-191	393	! Q9GL61 erythrogaus cuniculus (r
sp_vertebrate:Q97379	+	1726.00	3004.38	1.3e-159	421	! Q97379 gallus gallus (chicken). e
sp_human:Q16433	+	1648.50	2871.29	4.3e-152	318	! Q16433 homo sapiens (human). en
sp_vertebrate:Q91548	+	1623.00	2823.60	1.5e-149	415	! Q91548 xenopus laevis (afrikan
sp_vertebrate:Q9DEC5	+	1295.50	2248.33	1.7e-117	413	! Q9DEC5 oryzias latipes (medaka
sp_vertebrate:Q973868	+	1289.00	2236.32	7.4e-117	436	! Q973868 coturnix coturnix (commo
sp_vertebrate:Q9DGM2	+	1260.50	2186.50	4.5e-114	426	! Q9DGM2 brachydanio rerio (zebra
sp_mammal:Q9N0W7	+	1246.00	2160.66	1.2e-112	441	! Q9N0W7 erythrogaus cuniculus (r
sp_human:Q9UD23	+	1235.00	2148.33	5.8e-112	442	! Q9UD23 homo sapiens (human). en
sp_mammal:Q9N8V25	+	1238.50	2147.65	6.4e-112	434	! Q9N8V25 canis familiaris (dog).
sp_mammal:Q28467	+	961.00	1669.99	6.4e-85	176	! Q28467 macaca fascicularis (cra
sp_human:Q16432	+	958.00	1664.06	1.3e-84	187	! Q16432 homo sapiens (human). en
sp_mammal:Q9OGMC5	+	709.00	1230.18	2.6e-60	135	! Q9OGMC5 ovis aries (sheep). end
sp_vertebrate:Q9QYB8	+	688.00	1190.91	3.3e-58	168	! Q9QYB8 brachydanio rerio (zebra
sp_mammal:Q9BDJ8	+	620.50	1073.35	1.3e-51	153	! Q9BDJ8 ovis aries (sheep). end
sp_human:Q43441	+	532.00	921.47	5.1e-43	110	! Q43441 homo sapiens (human). en
sp_vertebrate:Q98TC7	+	480.00	825.13	7.4e-38	174	! Q98TC7 oryzias latipes (medaka
sp_rodent:Q9QW13	+	464.00	788.41	3.7e-36	384	! Q9QW13 rattus norvegicus (rat)
sp_rodent:Q9QRC6	+	438.50	738.70	1.4e-33	603	! Q9QRC6 ratuus sp. norvegicus (rat)
sp_rodent:Q9ZG03	+	414.50	696.60	3.1e-31	600	! Q9ZG03 mus musculus (mouse). g
sp_rodent:Q9QV42	+	414.50	696.60	3.1e-31	600	! Q9QV42 mus musculus (mouse). g
sp_mammal:Q95KN4	+	367.50	633.42	6.1e-27	101	! Q95KN4 erythrogaus cuniculus (r
sp_rodent:Q9QYC5	+	364.00	610.29	2.5e-26	481	! Q9QYC5 rattus norvegicus (rat)
sp_rodent:Q9QYV2	+	367.50	584.82	9.0e-25	348	! Q9QYV2 mus musculus (mouse). g
sp_rodent:Q98313	+	325.50	543.67	1.4e-22	438	! Q98313 mus musculus (mouse). m
sp_rodent:Q9R0D1	+	322.50	539.58	2.6e-22	393	! Q9R0D1 mus musculus (mouse). m
sp_rodent:Q9R1M0	+	322.50	539.58	2.6e-22	401	! Q9R1M0 mus musculus (mouse). m
sp_rodent:Q9R1L9	+	322.50	539.36	2.7e-22	401	! Q9R1L9 mus musculus (mouse). m
sp_rodent:Q9J1Y1	+	322.50	538.26	2.8e-22	444	! Q9J1Y1 mus musculus (mouse). m
sp_mammal:Q9S854	+	319.50	534.12	5.2e-22	400	! Q9S854 macaca fascicularis (cra
sp_rodent:Q95040	+	318.50	531.35	7.4e-22	405	! Q95040 mus musculus (mouse). m
sp_human:Q9H573	+	317.50	529.23	8.6e-22	454	! Q9H573 homo sapiens (human). d
sp_invertebrate:Q9NFV2	+	316.50	529.41	1.0e-21	380	! Q9NFV2 lymnaea stagnalis (gre
sp_mammal:Q9MW9	+	316.50	528.85	1.0e-21	400	! Q9MW9 macaca mulatta (rhesus
sp_invertebrate:Q9NFV0	+	314.50	525.90	1.6e-21	380	! Q9NFV0 lymnaea stagnalis (gre
sp_invertebrate:Q42324	+	313.50	524.05	2.0e-21	383	! Q42324 catostomus commersoni (W
sp_vertebrate:Q98U14	+	311.00	519.83	3.5e-21	377	! Q98U14 brachydanio rerio (zebra
sp_invertebrate:Q9NFV1	+	309.00	516.23	5.5e-21	380	! Q9NFV1 lymnaea stagnalis (gre
sp_invertebrate:Q9GVX8	+	308.50	512.28	6.8e-21	504	! Q9GVX8 drosophila melanogaste

```
885 TCCTATCAATGATTTAACTGCTGCTGGCGCTGGCCTTTTGATCAC 934
|||||
134 euProileAsnValPheLysLeuAlaGlyArgTrpPheGluGln 150
|||||
935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTTTTGAGAGT 984
|||||
151 AsnAspPheGlyValPheLysCysLysLeuPheProPheLeuGlnLys 167
|||||
985 CTCGGTGGGATCACCCTCTCTCAACCTCTCGCTCTTAGTGTGACAGGT 1034
|||||
167 rSerValGlyLeuThrValLeuAsnLeuCysAlaLeuSerValAspArg 184
|||||
1035 ACAGACAGTGGCTCTCTCGAGTGGTGTTCAGGGAATGGAATCCCTTG 1084
|||||
184 yrArgAlaValAlaSerTrpSerArgValGlnGlyLeuProLeu 200
|||||
1085 GTAACGTGCCATCAAAATGCTCCATCTGATCTGCTCTTATCTCTGGC 1134
|||||
201 ValThrAlaIleGluLeuValSerIleTrpIleLeuSerPheIleLeuAl 217
|||||
1135 CATTCCTGAAGCATGGCTCTCTCATCTGTGTCACCTTTGAATATAGGGGTG 1184
|||||
217 alieProGluAlaIleGlyPheValMetValProPheGluTrpLysGlyA 234
|||||
1185 AACAGCATAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
|||||
234 laGlnHisArgThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
|||||
1235 TACCAGATGTAAGACATGGTGGCTCTCTCGGCTTCTATTTCTGTATGCC 1284
|||||
251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTrpPheCysMetPr 267
|||||
1285 CTTGGTGTGCTACATCTCTTACACCTCTCATGACTGTGAGATGTGA 1334
|||||
267 oLeuValCysThrAlaIlePheTrpThrLeuMetThrCysGluMetLeuA 284
|||||
1335 ACAGAAGGAATGGCAGCTGTGAGAAATGCCCTCAGTGAACATCTTAAGCAG 1384
|||||
284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
|||||
1385 CGTCGAGAGTGCAGAAACAGTTTCTGCTTGGTGTAAATTTTGCTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 317
|||||
1435 TTGCTGTTTCCCTCTTCACTTAAGCCGTATATTGAAGAAACCTGTGATA 1484
|||||
317 ucYstrPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
|||||
1485 ACGAAATGGACAGAACCGATGGAATTAATTAAGTCTTCTTACGTCTATG 1534
|||||
334 spGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
|||||
1535 GATTACATCGGTATTAACCTGGCAACCATGAATTCATGATATAACCCCAT 1584
|||||
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProil 367
|||||
1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAATAATTTGTTCCAGTCTATGCC 1634
|||||
367 eaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
|||||
1635 TCTGCTGCTGCTGTGTACAGTCCAAAAGTCTGATGACCTCGGTCCTCATG 1684
|||||
384 euCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValProMet 400
|||||
1685 AACGGNACAGCATCCAGTGGGAGACACCATCAATTCATGATATAACCAAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnProGluGlnAsnAsnHisAsnth 417
|||||
1735 AGACCGGACGACCATTAAGCAGCATGAAC 1765
|||||
417 rGluArgSerSerHisLysAspSerIleAsn 427
|||||
seq_name: sp_rodent:Q91VV2
```

```
seq_documentation_block:
ID Q91VV2 PRELIMINARY; PRT; 427 AA.
AC Q91VV2;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE SIMILAR TO ENDOTHELIN RECEPTOR TYPE A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008277; AAH08277.1; -.
KW Receptor.
SQ SEQUENCE 427 AA; 48578 MW; 10142BE73B028233 CRC64;

alignment_scores:
Quality: 2126.00 Length: 427
Ratio: 5.086 Gaps: 0
Percent Similarity: 97.892 Percent Identity: 92.272

alignment_block:
US-09-931-157-1 x Q91VV2 ..
Align seg 1/1 to: Q91VV2 from: 1 to: 427

485 ATGGAACCCCTTTGCTCAGGCACTCTTTGGCTGGCAGCTGGTGGATG 534
|||||
1 MeSerIlePheCysLeuAlaIleTyrPheTrpLeuTrpMetValGlyG 17
|||||
535 TGTAACTCAGTGAATATCTCGAGATACAGCACAAATCTAAGCAATCATG 584
|||||
17 yValMetAlaAspAsnProGluArgTyrSerAlaAsnLeuSerHisM 34
|||||
585 TGGATGATTTTCCAGCTTTTTCGGTGGCAGACAGCTCAGCTTCTCTGTTACC 634
|||||
34 etGluAspPheThrProPheProGlyThrGluIleAsnPheLeuGlyThr 50
|||||
635 ACTCATCAACCCACTAATTTGGTCTTACCAGCAATGCTCAATGCACAA 684
|||||
51 ThrHisArgProProAsnLeuAlaLeuProSerAsnGlySerMetHisG 67
|||||
685 CTATTGCCCCACAGCAGACTAAATTTACTTCAGCTTTTCAAAATACATTAACA 734
|||||
67 yTyrCysProGlnGlnThrLysIleThrThrAlaPheLysTyrIleAsn 84
|||||
735 CTGTGATATCTTGTACTATTATTTCATCTGGGAATGGTGGGAATGCAACT 784
|||||
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
|||||
785 CTGCTCAGGATCATTTTACCAGAAACAATGATGAGGAATGCCCAACGC 834
|||||
101 LeuLeuArgIleIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
|||||
835 GCTGATAGCCAGTCTTGGCCCTTGGACACCTTATCTATGTGCTCATTTGATC 884
|||||
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValIleAspL 134
|||||
885 TCCTCATCAATGATTTTAACTGCTGCTGGCGCTGGCCTTTTGATCAC 934
|||||
134 euProileAsnValPheLysLeuAlaGlyArgTrpPheAspHis 150
|||||
935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTTTTGAGAGT 984
|||||
151 AsnAspPheGlyValPheLysCysLysLeuPheProPheLeuGlnLys 167
|||||
985 CTCGGTGGGATCACCCTCTCTCAACCTCTCGCTCTTAGTGTGACAGGT 1034
```

```
|||||
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
1035 ACAGAGCAGTGGCTCTCGAGTCTGTTCAGGGAATTGGGATTCCTTTG 1084
|||||
184 yArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
1085 GTAACGCCATTGAAATGCTCCATCGGATCGCTGCTCTTATCTGTCG 1134
|||||
201 IleThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeu 217
1135 CATTCTCAACGAGTGGCTCGTCATGGTACCCCTTTGCAATATAGGGGT 1184
|||||
217 aileProGluAlaIleGlyPheValMetValProPheGluTrpLysGly 234
1185 AACAGCAATAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTC 1234
|||||
234 luLeuHisArgThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
1235 TACCAAGATGTAAGGACTGGTGGCTCTTCGGGTCTATTCTCTATGCC 1284
|||||
251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTrpPheCysMetPr 267
1285 CTTGGTGTGCACTCGGATCTTACACCCCTCATCTGTGAGATGTGA 1334
|||||
267 oLeuValCysThrAlaIlePheTrpThrLeuMetThrCysGluMetLeu 284
1335 ACAGAAGTAATGGCAGCTTGAGATTCGCCCTCAGTGAACATCTTAAGCAG 1384
|||||
284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
1385 CGTCGCAAGTGGCAAAACAGTTTCTCGTGTGTTGTTGTAATTTTGTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 317
1435 TTGCTGTCTCCTCTTACATTAAGCCGTATATTCAAGAAACTGTGTATA 1484
|||||
317 ucYsrPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 ACGAAATGGCAAGACCGATGTAATTAATTAATTAATTAATTAATTAAT 1534
|||||
334 spGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
1535 GATTACATCGGTATTACTTGGCAACCATGAATTCATGTATTAACCCCAT 1584
|||||
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
1585 AGCTCTGTATTCTGACGACGAATTTAAATTTGTTTCCAGTCATGCC 1634
|||||
367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TCTGCTGTGCTGTATACAGTCCAAAAGTCTGATGACCTCGCTGCCCATG 1684
|||||
384 euCysCysCysCysHisGlnSerLysSerLeuMetThrSerValProMet 400
1685 AACGGACAACATCCAGTGGGAAGACACGATCAAAACACCAACACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnGlnGlnAsnAsnHisAsnTh 417
1735 AGACCGGACGACCATAGGACACATCAAC 1765
|||||
417 rGluArgSerSerHisLysAspSerMetAsn 427
seq_name: sp_mammal:Q9GL61
```

seq_documentation_block:

ID Q9GL61 PRELIMINARY; PRT; 393 AA.

AC Q9GL61;

DT 01-MAR-2001 (TrEMBLrel. 16, Created)

DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)

DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)

DE ENDOTHELIN RECEPTOR TYPE A (FRAGMENT).

OS Oryctolagus cuniculus (Rabbit).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RA Teupser D., Wilfert W., Thierry J.;
RL Submitted (Oct-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF311974; AAG31802.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1.
DR PRINTS: PR00237; GPCR_RHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1
FT NON_TER 393
SQ SEQUENCE 393 AA; 44903 MW; B121FD553835EF02 CRC64;

alignment_scores:

Quality: 2046.00 Length: 393
Ratio: 5.233 Gaps: 0

Percent Similarity: 99.491 Percent Identity: 97.201

alignment_block:

US-09-931-157-1 x Q9GL61

Align seg 1/1 to: Q9GL61 from: 1 to: 393

```
536 GTAATCAGTGATAATCTCGAGAGATACAGACAAATCTAAGCAATCATGT 585
|||||
1 ValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisMe 17
586 GGATGATTTCACCACTTTCTCGTCAGACAGCTCAGCTTCTGTTACCA 635
|||||
17 tAspGluPheThrThrPheHisGlyProGluLeuAsnLeuValThrT 34
636 CTCATCAACCCCAATTTGCTCTACCCAGCAATGGCTCAATGCACAAC 685
|||||
34 hrHisArgProThrAsnLeuValLeuProSerAsnGlySerArgHisAsn 50
686 TATTGCCACACAGACACTAAATTAATTAATTAATTAATTAATTAATTA 735
|||||
51 TyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnTh 67
736 TGTGATATCTTGTTACTATTTTCATCGTGGGAATGTGGGGAATGCAACT 785
|||||
67 rValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThrL 84
786 TGCTCAGGATCATTTACAGAACAAATGTATGAGGAATGGCCCCAACGCG 835
|||||
84 euLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAla 100
836 CTGATAGCCAGTCTGCCCTTGGAGACCTTATCTATGTGGTCATGTATCT 885
|||||
101 LeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspLe 117
886 CCCTATCAATGTTTAAAGCTGCTGGTGGGCTGGGCTGCTTTTGTATCACA 935
|||||
117 uProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHisA 134
936 ATGACTTTGGCGTATTTCTTTTCAAGCTGTTCCTCCCTTTTTCGAGAAGTCC 985
|||||
134 snAspPheGlyValPheLeuLysLysLeuPheProPheLeuGlnLysSer 150
986 TCGGTGGGGATCACCGCTCCTCAACCTCTGCCCTCTTAGTGTGTGACAGTA 1035
|||||
151 SerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTy 167
1036 CAGAGCAGTTCCTCCCTGGAGTCTGTGTTCAGGGAATGGGATTCCTTTGG 1085
|||||
167 rArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeuI 184
```

```

1086 TAACCTGCCATGAATTTGCTCCATCTGGATCCTGCTCTTATCTCTGGCC 1135
      :|||||
184  lethrAlaileGluIleValserIleTrpIleLeuSerPheIleLeuAla 200
      :|||||
1136 ATTCTCAAGCGATTGCTCGTCATGGTACCCCTTTGAATATAGGGTGA 1185
      :|||||
201  IleProGluAlaIleGlyPheValMetValProPheGluTrpArgGlyG 217
      :|||||
1186 ACAGCAATAAACCTGTATGCTCAATGCCACATCAAAATTCATGAGTTCT 1235
      :|||||
217  uGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPheT 234
      :|||||
1236 ACCAAGATGAAGACTGTGGCTCTCGGGTTCTATTCTGTATGCC 1285
      :|||||
234  yrGlnAspValLysAspTrpIlePheGlyPheTyrPheCysMetPro 250
      :|||||
1286 TTGGTGTGCACTCGGACTCTCTACACCTCATGACTTGTGAGATGTGAA 1335
      :|||||
251  LeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuAs 267
      :|||||
1336 CAGAAGGAATGGCAGCTTGAGATTGCCCTCAGTGAACATCTTAAGCAGC 1385
      :|||||
267  nArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGlnA 284
      :|||||
1386 GTCGAGAAAGTGGCAAAACAGTTTCTGCTGTGGTGTAAATTTTGCCTT 1435
      :|||||
284  rArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLeu 300
      :|||||
1436 TGCTGTCTCCTCTCTCACTTAAGCCGTATATTGAAGAAACTGTGTATAA 1485
      :|||||
301  CysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrAs 317
      :|||||
1486 CGAAATGGCAAGAACGGATGTGAATTTACTTAGTTCTTCTGCTCATGG 1535
      :|||||
317  pGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMetA 334
      :|||||
1536 ATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAAACCACATA 1585
      :|||||
334  spTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIle 350
      :|||||
1586 GCTCTGTATTTTGTGAGCAAGAAATTTAAATTTTCCAGTCATGCCCT 1635
      :|||||
351  AlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysLe 367
      :|||||
1636 CTGCTGCTGCTGTATACCAGTCCAAAGCTCTGATGACCTCGCTCCCATGA 1685
      :|||||
367  uCysCysCysCysHisGlnSerLysSerLeuMetThrSerValPrometa 384
      :|||||
1686 ACGGAACAAGCATCCAGTGGGAAGAACAC 1714
      :|||||
384  snGlyThrSerIleGlnTrpLysAsnHis 393
      :|||||

```

seq_name: sp_vertibrate:073739

seq_documentation_block:

```

ID 073739 PRELIMINARY; PRT; 421 AA.
AC 073739;
DT 01-AUG-1998 (TReMBLrel. 07, Created)
DT 01-AUG-1998 (TReMBLrel. 07, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ENDOTHELIN TYPE A RECEPTOR.
GN EDNRA.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
ON NCBI_TaxID=9031;
RX STRAIN=WHITE LEGHORN;
RA MEDLINE=99030274; PubMed=9811577;
RT "Pharmacological inactivation of the endothelin type A receptor in the

```

```

RT early chick embryo: a model of mispatterning of the branchial arch
RL derivatives";
RL Development 125:4931-4941(1998).
DR EMBL: AF040634; AAC77793.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PRINTS: PS00237; GPCRHHODPSN.
DR PROSITE: PS00237; G_PROTEIN_RECEP_FL_1; UNKNOWN_1.
DR PROSITE: PS50262; G_PROTEIN_RECEP_FL_2; 1.
KW Receptor..
SQ SEQUENCE 421 AA; 47713 MW; 33A2C97A6DCD02DF CRC64;

```

alignment_scores:

```

Quality: 1726.00 Length: 427
Ratio: 4.530 Gaps: 2
Percent Similarity: 89.227 Percent identity: 77.283

```

alignment_block:

US-09-931-157-1 x 073739

Align seg 1/1 to: 073739 from: 1 to: 421

```

485 ATGGAACCCCTTTCCTCAGGCGCATCTTTGGTGGCAGCTGGTTGGATG 534
      |||||
1 MetGluAlaLeuTyrLeuArgValSerSerLeuLeuLeuLeuGlyPh 17
      :|||||
535 TGTATCAGTGTATATCTCAGAGATACAGCACAAATCTAAGCAATCATG 584
      :|||||
17 eValLeuCysAspSerSerAspLysTyrAlaThrAsnTrpSerAspValG 34
      :|||||
585 TGGATGATTTCCACCACCTTTTCGTGCGCAGAGCTCAGCTTCTGCTTACC 634
      :|||||
34 lylleTyrLeuProThrSerSerGlyIleGluSerSerLeuLeuProThr 50
      :|||||
635 ACTCATCAACCCACTAATTTGGTCTCTACCCAGCAATGCTCAATGCACAA 684
      :|||||
51 ThrArgArgProPro.....ValSerAsnGlnThrValLys... 62
      :|||||
685 CTATTGCCCCACAGCAGACTAAATTTACTTCAGCTTTCAAATACATTAACA 734
      :|||||
63 ...CysSerGlnGlnThrLysIleAlaGluThrPheLysTyrIleAsnTr 78
      :|||||
735 CTGTGATATCTGTACTATTTCATCTCGTGGAAATGCTGGGAATGCAACT 784
      :|||||
78 hrValValSerCysAlaIlePheIleValGlyMetValGlyAsnAlaThr 94
      :|||||
785 CTGCTCAGGATCATTTACCAGACAAATGTATGAGGAATGCCCAACGC 834
      :|||||
95 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 111
      :|||||
835 GCTGATAGCCAGTCTTGCCTTGGAGACCTTATCTATGTGTGTCATTGATC 884
      :|||||
111 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrIleValIleAspI 128
      :|||||
885 TCCCTATCAATGTATTTAAGCTGCTGGTGGCGCTGCCCTTTTCATCAC 934
      :|||||
128 leProIleIleValTyrLysLeuLeuAlaGlnLysTrpProPheGlyAsp 144
      :|||||
935 AATGACTTTGGCGTATTTCTTTCGAAGCTGTTCCTCTTTTTCGAGAACTC 984
      :|||||
145 SerGluPheGlyGlnPheLeuCysLysPheLeuProPheIleGlnLysAl 161
      :|||||
985 CTGCGTGGGATCACCCTCTCAACCTCTGCGCTCTTAGTGTGTGACAGGT 1034
      :|||||
161 aSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 178
      :|||||
1035 ACAGACAGCTTGCCTCTCTGGAGTGTGTTCAGGGAATGGGATTCCTTTG 1084
      :|||||
178 yrArgAlaValAlaSerTrpSerArgValGlnGlnIleGlyIleProMet 194
      :|||||
1085 GTAACTCCCATTTGAAATTTGCTCCATCTGATCTGCTCTTATCTTCTGGC 1134
      :|||||

```

```
195 IleThrAlaIleGluIlePheSerIleTrpLeuLeuSerPheIleLeuAl 211
1135 CATTCCTGAAGCGATTGGCTTGGTCATGTAACCTTTTGAATAGGGGTG 1184
DR EMBL; S81545; AAB36327.1; -
DR InterPro; IPR00276; GPCR_Rhodpsn.
211 atleProGluAlaIleGlyPheAlaValProPheArgTyrLysAspG 228
DR pfam; PF00001; 7tm.1; 2.
DR PROSITE; PS0262; G_PROTEIN_FCEP_FL_2; 2.
1185 AACAGCATAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
KW Receptor.
228 luserTyrValThrCysMetLeuAsnProThrAsnLysPheMetLeuPhe 244
alignment_scores:
1235 TACCAAGATGTAAGGACTGGTGGCTCTTCGGGTTCTATTTCGTATGCC 1284
Quality: 1648.50 Length: 427
Ratio: 5.184 Gaps: 1
245 TyrLysAspAlaLysAspTrpLeuPheGlyPheTyrPheCysMetPr 261
Percent Similarity: 74.473 Percent Identity: 74.473
1285 CTTGGTGTCCACTGGCATCTTACACCTCATGACTTGTGAGATTTGA 1334
alignment_block:
261 oLeuAlaCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 278
US-09-931-157-1 x Q16433 ..
1335 ACAGAAGGAATGGCAGCTTGAGAATGGCCCTCAGTCAACATCTTAAGCAG 1384
Align seg 1/1 to: Q16433 from: 1 to: 318
278 snArgArgAsnSerAsnLeuArgIleAlaLeuSerGluHisLeuLysGln 294
1385 CGTCGAGAAGTGGCAAAACAGTTTCTGCTGGTGTAAATTTTGTCTCT 1434
295 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 311
1435 TTGCTGGTTCCTCTTCACTTAAGCGTATATTGAAGAACTGTGTATA 1484
311 ucysTrpPheProLeuHisLeuSerArgIleLeuLysLysMetValTyrA 328
1485 ACGAATGACAGAAGACCATGTAATTAATTAATTAATTAATTAATTAAT 1534
328 snGluArgAspProGlyArgCysGluLeuLeuSerPheLeuLeuProLeu 344
1535 GATTACATCGGTATTAACTTGGCAACCATCAATTCATGTATAAACCCTAT 1584
345 AspTyrIleSerIleAsnLeuAlaThrMetAsnSerCysIleAsnProI 361
1585 AGCTGTATTTGTGAGCAAGAAATTTAAATTTGTTCCAGTCATGCC 1634
361 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 378
1635 TCTGCTGCTGCTTACCAGTCAAAAGTCTGATGACCTCGGTCCTCCATG 1684
378 euCysCysCysSerGlnSerLysSerLeuAlaThrSerValProMet 394
1685 AACGGAAACAGATCCAGTGGAGAACACACGATCAAAACAAACACACAC 1734
395 AsnGlyThrSerIleGlnTrpLysAsnGlnGluLeuAsnAsnHisAsnTh 411
1735 AGACGGAGCAGCCATAAGACAGCATGAAC 1765
411 rAspArgSerSerHisLysAspSerIleAsn 421
```

seq_name: sp_human:Q16433

seq_documentation_block:

```
ID Q16433 PRELIMINARY; PRT; 318 AA.
AC Q16433;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE ENDOTHELIN-A RECEPTOR DELTA 3-4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96190719; PubMed=8611157;
RA Miyamoto Y., Yoshimasa T., Arai H., Takaya K., Ogawa Y., Itoh H.,
RA Nakao K.;
RT "Alternative RNA splicing of the human endothelin-A receptor generates
```

```
RT multiple transcripts.";
RL Biochem. J. 313:795-801(1996).
DR EMBL; S81545; AAB36327.1; -
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR pfam; PF00001; 7tm.1; 2.
DR PROSITE; PS0262; G_PROTEIN_FCEP_FL_2; 2.
KW Receptor.
SQ SEQUENCE 318 AA; 36488 MW; 0423FAA16DB6A9BF CRC64;

alignment_scores:
Quality: 1648.50 Length: 427
Ratio: 5.184 Gaps: 1
Percent Similarity: 74.473 Percent Identity: 74.473
alignment_block:
US-09-931-157-1 x Q16433 ..
Align seg 1/1 to: Q16433 from: 1 to: 318
485 ATGGAACCCCTTTGGCTCAGGGCATCCTTTTGGCTGGCACTGGTTGGATG 534
1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyCy 17
535 TGTAAATCAGTGATAATCCTGAGAGATACACACAAATCTAAGCAATCATG 584
17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisV 34
585 TGGATGATTTACACACTTTTTCGTGGCAGAGCTCAGCTTCCTGGTTACC 634
34 aAspAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
635 ACTCATCAACCCACTAATTTGGTCTTACCAGCAATGGCTCAATGCACAA 684
51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMethHisAs 67
685 CTATTGCCACACAGACATAAAATTTACTTCAGCTTTCAAAATACATAACA 734
67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
735 CTCTGATATCTTGTACTATTTTTCATCGTGGGAATGGTGGGAATGCAACT 784
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
785 CTCTCAGGATCATTTACCAGAACAAATGTATGAGGAATGCCCAACGC 834
101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
835 GCTGATAGCCAGCTCTTCCCTTGGAGACCTTATCTATGTGTCATTCATC 884
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValIleAspL 134
885 TCCCTATCAATGTATTTAAGCTGTGCTGGCGCTGGCCTTTTGTATC 934
134 euProIleAsnValPheLys..... 140
935 AATGACTTTGGCGGTATTTCTTTCAGAGCTGTTCCTCCCTTTTTCGAGAAGTC 984
140 ..... 140
985 CTGGTGGGGATCACCGCTCCTCAACCTCTCGGCTCTTAGTGTGCACAGGT 1034
140 ..... 140
1035 ACAGAGCAGTTGCTCTCTGGAGTCGTGTTCCAGGAATTTGGGATTCCTTTG 1084
140 ..... 140
1085 GTAACCTGCCATTGAAATTTGCTCCATCTGGATCCTGCTTTTATCTTGGC 1134
140 ..... 140
1135 CATTCCTGAAGCGATTGGCTTGGTCATGTAACCTTTTGAATAGGGGTG 1184
```

140 140

1185 AACAGCATAAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
141Phe 141

1235 TACCAAGATGTAAGGACTGTGGCTCTTCGGTTCATTCTGTATGCC 1284
142 TyrGlnAspValLysAspTrpLeuPheGlyPheYrPheCysMetPr 158

1285 CTGGTGTGCACCTCGGATCTCTACACCCCTCATGCTGTGAGATGTTGA 1334
158 oleuValCysThrAlaIlePheYrThrLeuMetThrCysGluMetLeuA 175

1335 ACAGAAGAAATGGCAGCTTGAGAAATGCCCTCACTGAACATCTTAAGCAG 1384
175 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 191

1385 CGTCGAGAAGTGGCAAAACAGTTCTTCGTGTTGGTTGTAATTTTGGTCT 1434
192 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 208

1435 TTGCTGTTCCCTCTTCCACTTAAGCCGTATATTGAAGAAACATGTGTATA 1484
208 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyra 225

1485 ACGAATGGACAAGAACCGATGTCGAATTAATTAGTTTCTTACTGCTCATG 1534
225 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 241.

1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAACCCCAT 1584
242 AspTyrlleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIle 258

1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTGTTCCAGTCATGCC 1634
258 eAlaLeuTyrrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 275

1635 TCTGCTGCTGTTTACCAGTCCAAAAGTCTGATGACCTCGGTCCCGCATG 1684
275 euCysCysCysCysTyrglnSerLysSerLeuMetThrSerValProMet 291

1685 AACGGAAACAAGCATCCAGTGGAGAACCACCGATCAAAACAAACCAACAC 1734
292 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnAsnHisAsnTh 308

1735 AGACCGAGCAGCCOATTAAGCAGCAGCATGAAC 1765

308 rAspArgSerSerHisLysAspSerMetAsn 318

seq_name: sp_vertebrate:091548

```

seq_documentation_block:
ID      Q91548      PRELIMINARY;      PRT;      415 AA.
AC      Q91548;
DT      01-NOV-1996 (TRENBLrel. 01, Created)
DT      01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE      ENDOTHELIN RECEPTOR.
OS      Xenopus laevis (African clawed frog).
OC      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC      Amphibia; Batrachia; Anura; Mesobatrachia; Pipelidae;
OC      Xenopodinae; Xenopus.
OX      NCBI_TaxId=8355;
RN      [1]
RS      SEQUENCE FROM N.A.
RC      TISSUE=HEART;
RX      MEDLINE=94230448; PubMed=8175772;
RA      Kumar C., Mwangi V., Nuthallaganti P., Wu H.L., Pullen M., Brun K.,
RA      Aiyar H., Morris R.A., Naughton R., Nambi P.;
RT      "Cloning and characterization of a novel endothelin receptor from
RT      xenopus heart.";

```

```

RL J. Biol. Chem. 269:13414-13420(1994).
DR EMBL: U06633; AAA19570.1; -.
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm1_1.
DR PRINTS: PR00237; GPCRHHODPSN.
DR PROSITE: PS00237; G-PROTEIN_RECEP_FL1; UNKNOWN_1.
DR PROSITE: PS0262; G-PROTEIN_RECEP_FL2; 1.
DR KW Receptor.
SQ SEQUENCE 415 AA; 47003 MW; 98DA919097085D70 CRC64;

alignment_scores:
    Quality: 1623.00
    Ratio: 4.340
    Percent Similarity: 88.626

alignment_block:
    US-09-931-157-1 x Q91548
    Length: 422
    Gaps: 4
    Percent Identity: 72.512

```

Align seg 1/1 to: Q91548 from: 1 to: 415

```

500 CTCAGGCAATCCTTTTGCTGGCACTGTGGTGGATGTGAATCATCGATGATAA 549
      ||||| ::: |||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
6   LeuArgPheThrValLeuLeuValLeuAlaGlyIleAlaValSerSerSe 22

550 TCCTGAGAGATACAGCAAATCTAAGCAATCATGTGGATGATTTCACCA 599
      : ||||| ||| :::::: |||||||||
22  rPheGlyGluTyrGlnAsnArgThrAspAlaSerThrAspPheThr. 38

600 CTTTTCGTGGCACAGACTGCAGTTCCTGGTTACCACCTCATCAACCCACT 649
      39 ..... MetLeuAsnArgSerHis.....Thr 45

650 AATTGTTGCTCTACCAGCAATGGCTCAATGCACAACATTGGCCACAGCA 699
      : ||| ||||| ||| ::||| |||||::|||
46 SerProValArgLysGlyAsnArgSerAla... AspLeuCysProGluLys 61

700 GACTAAATTAATTCACGCTTTCAATACATTAACACTGTGATATCTTTGTA 749
      : |||||::: |||||::|||::|||::|||::|||::|||::|||::|||::|||
61 sThrLysIleAsnHisValPheLysTyrIleAsnThrIleLeuSerCyst 78

750 CTATTTTCATCGTGGGAATGTGGGAATGCAACTCTGCTCAGGATCATTT 799
      |||||PhelIleIleGlyMetValGlyAsnAlaThrLeuLeuArgIleIle 94

800 TACCAGAACAATGTATGAGGAATGGCCCCAACGGCTGATGCCAGTCT 849
      95 TyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeuIleAlaSerLe 111

850 TGCCCTTGGAGACCTTATCTATGTGTGCTATTCATCTCCCTATCAATGTA 899
      ||||| ||||| |||||::|||::|||::|||::|||::|||::|||::|||
111 uAlaLeuGlyAspLeuIleTyrIleValIleAspIleProIleIleValP 128

900 TTAAAGTCGTGGCTGGCGCTGGCGCTTTTGATPCAAATCACTTTGGCGTA 949
      |||||::: |||||::|||::|||::|||::|||::|||::|||::|||
128 helysLeuLeuGlyGlnArgTrpPheAspGlnSerProValGlyAla 144

950 TTTCTTTGGAAGCTGTTCCTCTTTTGAGAGAGTCTCTCGGTGGGGATCAC 999
      145 PheLeuCysLysLeuValProPheIleGlnLysAlaSerValcglyIleTh 161

1000 CGTCCTCAACCTCTGCCTCTTAGTGTGTACAGGTPACAGACGAGTTCGCT 1049
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
161 rValLeuAsnLeuCysAlaLeuSerValAspArgTyrArgAlaValProS 178

1050 CCTGGAGTCGTGTACAGGAATTCGGGATTCCTTTGGTAACTGCATTTGAA 1099
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
178 erTrpSerArgValGlnGlySerGlyIleProLeuIleThrAlaIleGlu 194

1100 ATTGCTCCNATCGGATCCTGTCTCTTATCTCTGGCCATTCCTGGAACGGAT 1149
      195 IleIleSerIleTrpValLeuSerPheValLeuAlaIleProGluAlaIl 211

```

```
1150 TGCTTCGTCATCGTACCCCTTTGAATATAGGGGTGAACACCAATAAACCT 1199
|||||
211 eGlyPheValMetValProPheGlyuTyArgGlyGluGlnPheArgThrC 228
|||||
1200 GTATGCTCAATGCCACATCAAAATTCATGGAGTCTACCAAGATGTAAG 1249
|||||
228 yMetPheHisAlaThrSerProPheMetMetPheTyLysAsnAlaLys 244
|||||
1250 GACTGTGGCTCTTCGGGTTCTATTCTGTATGCCCTTGCTGGCTACTGC 1299
|||||
245 AspTrpTrpLeuPheGlyLeuTyPheCysValProLeuAlaCysThrGl 261
|||||
1300 GATCTTACACCTCATGACTTGTGAGATGTTGAACAGGAAGTAAGCA 1349
|||||
261 yValPheTyThrMetMetThrCysGluMetLeuHisGlnArgLysGlyS 278
|||||
1350 GCTTGAGAATGGCCCTCAGTGACATCTTAAGCAGCGCTGAGAGTGGCA 1399
|||||
278 erLeuArgIleAlaLeuSerGluHisLeuLysGlnArgArgGluValAla 294
|||||
1400 AAACAGTTTTCTGCTTGGTTGTAATTTTGGCTCTTTGGCTGGTTCCCTCT 1449
|||||
295 LysThrValPheCysLeuValValIlePheAlaLeuCysTrpPheProLe 311
|||||
1450 TCACCTTAAGCGGTATATTGAAGAAACTCTGTATAACGAATGCACAGA 1499
|||||
311 uHisLeuSerArgIleIleLysAsnThrIleTyAsnGluLeuAspProS 328
|||||
1500 ACCGATCGAATTACTTAGTTTCTTACTGCTCATGATTACATCGGTATT 1549
|||||
328 erArgCysGluLeuLeuSerPheLeuLeuValMetAspPheIleSerIle 344
|||||
1550 AACTTGGCAACCATGAATTCATGTATAAACCCCATAGCTCTGTATTGTT 1599
|||||
345 AsnLeuAlaAlaLeuAsnSerCysIleAsnProIleAlaLeuTyPheVa 361
|||||
1600 GAGCAAGAAATTTAAATTTGTTTCCAGTCATGCTGCTGCTGCTGCTGTT 1649
|||||
361 lSerLysLysPheLysAsnCysPheGlnSerCysLeuCysCysCys. 377
|||||
1650 ACCAGTCCAAAGTCTCATGACCTCGGTGCCCATGAACGAAACAAAGATC 1699
|||||
378 ..GlnSerLysThrHisIleAsnThrAlaProMetAsnValThrSerIle 393
|||||
1700 CAGTGGGAAGACCAACGATCAAAACCAACCAACACGACGCGGACGCCA 1749
|||||
394 GlnTrpLysAsnHisAspGlnAsnTyTyArgAlaAspArgSerIleHi 410
|||||
1750 TAAGGACAGCATGAAC 1765
|||||
410 sLysAspSerIleAsn 415
```

seq_name: sp_vertibrate:Q9DEC5

```
seq_documentation_block:
ID Q9DEC5 PRELIMINARY; PRT: 413 AA.
AC Q9DEC5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENDOTHELIN RECEPTOR.
GN EDNR.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tobica-Teramato T., Akiyama T.;
RT "Endothelin Receptor in Medaka Melanophores. I. Cloning and its.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
```

```
DR EMBL; AB045356; BAB20097.1; ..
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 413 AA; 46752 MW; 92E99424CEBA854C CRC64;

alignment_scores:
  Quality: 1295.50      Length: 381
  Ratio: 3.962         Gaps: 6
  Percent Similarity: 85.827      Percent Identity: 62.730

alignment_block:
US-09-931-157-1 x Q9DEC5 ..
Align seq 1/1 to: Q9DEC5 from: 1 to: 413

635 ACTCATCAACCCACTAATTTGGTCTCTACCCAGCAATGGCTCAATGCACAA 684
|||||
37 ThrAsnSerSerValGlnLeuProLeuProGlnProGlyArgProArgSe 53
685 C.....TATTGCCACACAGCAGACTAAATTTACTTTCAGCTTTCA 722
|||||
53 iProProProProMetCysValLysProThrGluIleLysAspAlaPheL 70
723 AATCATTAACACTGTGATATCTTGTTACTATTTTCATCGTGGGAATGGT 772
|||||
70 yTyTrValAsnThrIleIleSerCysLeuIlePheValValGlyIleIle 86
773 GGGNAATCAACTCTGCTCAGGATCATTTACCAGAACAAATGTATCAGAA 822
|||||
87 GlyAsnSerThrLeuLeuArgIleIleTyArgAsnLysCysMetArgAs 103
823 TGGCCCCCAACGCTGATAGCCAGCTTGGCCCTTGGAGACCTTATCTATG 872
|||||
103 nGlyProAsnValLeuIleGlySerLeuAlaLeuGlyAspLeuLeuTyI 120
873 TGGTCATTGATCTCCCTATCAATGATTTAAAGCTGCTGGCTGGCGGTGG 922
|||||
120 leIleIleAlaIleProIleAsnValTyLysLeuIleAlaGluAspTrp 136
923 CTTTTGATCACAATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTTT 972
|||||
137 Pro.....PheGlyValTyIleCysLysLeuMetProPh 148
973 TTTCGAGAAGTCTCTCGGTGGGATCACCGCTCTCAACCTCTCGCTCTTA 1022
|||||
148 eIleGlnLysAlaSerValGlyIleThrValLeuSerLeuCysAlaLeuS 165
1023 GTGTTGACAGGTACAGACAGCTTGCCTCCTGGAGTCGTTGTTTCAGGGA 1072
|||||
165 erIleAspArgTyHisAlaValThrSerTrpSerArgValLysGlyMet 181
1073 GGGATTCCTTTGGTAACCTGCCATTGAAATTTCTCCATCTCGATCCTGTC 1122
|||||
182 GlyIleProLeuTrpLysAlaValGluValThrLeuIleTrpLeuValAl 198
1123 CTTTATCTGCGCATTCCTGAAGCGATTTGGCTTCGTCATGCTACCTTTG 1172
|||||
198 aValValLeuAlaValProGluAlaLeuAlaPheAspMetLeuGluMet 215
1173 AATATAGGGGTGAACAGCATATAAACCTGTATGCTCAAT.....GCCACA 1216
|||||
215 roTyArgGlyAsnLysLeuArgIleCysLeuLeuHisProGluGlnPro 231
1217 TCAAAATTCATGAGTCTTACCAAGATGTAAGGACTGGTGGCTCTTCGG 1266
|||||
232 ThrValTyMetLysPheTyGlnGluAlaLysAspTrpTrpLeuPheGl 248
1267 GTTCTATTCTGTATGCCCTTGGTGTGCTGACATGCGGATCTTCTACACCC 1316
```



```
|||||
248 yPheTyrPheCysLeuProLeuAlaCysThrGlyIlePheTyrThrLeuM 265
1317 TGACTTGTGAGATTGAACAGAGGAATGCGAGCTTGAGAAATGCCCTC 1366
265 etThrCysGluMetLeuSerArgLysLysGly...MetArgIleAlaLeu 280
1367 AGTGAACATCTTACGACGCTGAGAGAGTGGCAAAACAGTTTCTGCTT 1416
281 AsnAspHisMetLysGlnArgArgGluValAlaLysThrValPheCysLe 297
1417 GGTGTAAATTTTCTGCTTCTGCTGCTGCTTCTGCTTCTGCTTCTGCT 1466
297 uValLeuIlePheAlaPheCysTrpLeuProLeuHisLeuSerArgIleL 314
1467 TGAAGAAACATGTGTATACGAATATGCAAGAACCGATGTGAATTAATT 1516
314 euLysLysThrValTyrAsnGluAsnAspProAsnArgCysGluLeuLeu 330
1517 AGTTTCTTACTGCTCATGATTACATCGGTATTAACCTGGCAACCATGAA 1566
331 SerPheLeuLeuValMetAspTyrIleGlyIleAsnMetAlaSerLeuAs 347
1567 TTCATGTATAAACCACCATAGCTCTGTATTTGTGACGAAAGAAATTTAAA 1616
347 nSerCysIleAsnProIleAlaLeuTyrPheValSerGlnLysPheLysA 364
1617 ATTGTTTCCAGTCATGCTCTGCTGCTGCTTACCAGTCCAAAAGTCTG 1666
364 snCysPheGlnSerCysLeuCysCysTrpCysTyrArgThrSerProLeu 380
1667 ATGACCTGGTCCCTCCATGAAGCAACCAAGCATCCAGTGGGAAGCAACCA 1716
381 .....AspGluArgGlySerGlyArgTrpLysGlySerCy 393
1717 TCAAAACACCAACACAGACGGGACCGACCCATAGGACAGC 1759
393 sGlnValAsn...GlyLeuAspArgThrSerSerArgSer 406
seq_name: sp-vertebrate:073868
seq_documentation_block:
ID 073868 PRELIMINARY; PRT; 436 AA.
AC 073868;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN RECEPTOR B SUBTYPE 2 (FRAGMENT).
GN EDNRB2
OS Coturnix coturnix (common quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=9091;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98169490; PubMed=9501209;
RA Lecoin L., Sakurai T., Ngo M.T., Yanagisawa M., Le Douarin N.M.;
RT "Cloning and characterization of a novel endothelin receptor subtype
in the avian class";
EL Proc. Natl. Acad. Sci. U.S.A. 95:3024-3029(1998).
DR EMBL; Y16089; CAA76055.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 2.
DR PRINTS; PR00237; GPCR_RHODPSN
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_FL1; UNKNOWN_1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_FL2; 1.
KW Receptor.
FT NON_TER 436 436
SQ SEQUENCE 436 AA; 49109 MW; 3FCCBE170B8FF22A CRC64;
alignment_scores:
Quality: 1289.00 Length: 408
Ratio: 3.942 Gaps: 7
Percent Similarity: 80.147 Percent Identity: 60.784
alignment_block:
US-09-931-157-1 x 073868
Align seg 1/1 to: 073868 from: 1 to: 436
599 ACTTTTCTGTCGACAGAGCTCAGCTTC...CTGGTTTACACTCATCAACC 645
|||||
28 ThrPheGlnValSerSerIleProPheGluAlaLeuSerGlnGluGlnAl 44
646 CACTAATTTGCTCTCTACCCAGC.....AATGGCTCAATGCACAAC 686
|||||
44 aTyrSerLeuValGlnProSerLeuPheGlnAspAlaLysAlaProAsn 61
687 AT..... 688
||
61 yrSerGluSerLeuProLysSerSerGlySerGluProProLeuLeuPro 77
689 ...TGCCACAGCAGACTAAATACCTTCAGCTTCAATACATTAACAC 735
|||||
78 ValCysAlaLysProAlaAspIleArgHisIlePheLysTyrIleAsnTh 94
736 TGTGATATCTTGTACTATTTTCATCGTGGGAATGTTGGGAATGCAATC 785
|||||
94 rIleValSerCysThrIlePheIleValGlyIleIleGlyAsnSerThrL 111
786 TGCTCAGGATCATTTACCAAGCAAAATGTATGAGGAATGGCCCAACGCG 835
|||||
111 euLeuArgIleIleTyrLysAsnLysCysMetArgAsnGlyProAsnVal 127
836 CTGATACCCAGCTTGGCTTGGAGACCTTATCATGTGCTGCTTATGATCT 885
|||||
128 LeuIleAlaSerLeuAlaLeuGlyAspLeuLeuTyrIleLeuIleAlaLe 144
886 CCCTATCAATGTATTAAAGCTGCTGGCGCTGGCGCTTTTGGATCACA 935
|||||
144 uProIleAsnValTyrLysLeuLeuAlaLysAspTrpPro..... 157
936 ATGACTTTGGCGTATTTCTTTCAGAGCTGTTCCTTTTTCAGAGATGCC 985
|||||
158 ....PheGlyValGlnValCysLysLeuValProPheIleGlnLysAla 172
986 TCGTGGGATCACCGCTCAACCTCTGCTGCTTGTAGTGTGACAGGTA 1035
|||||
173 SerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTy 189
1036 CAGAGCAGTTGCTCTGAGTGTTCAGGGAATGGGATTCCTCTTGG 1085
|||||
189 rArgAlaValAlaSerTrpSerArgIleGlnGlyIleGlyIlePromet 206
1086 TAACTGCCATTGAATTTGCTCCTCATCTGGATCTCTCTTATCTCTGCC 1135
|||||
206 rLysAlaValGluValThrLeuIleTrpAlaValAlaIleValLeuAla 222
1136 ATTCTCGAACCGATTGCTTCGTCATGCTACCTCTTGAATATAGGGGTGA 1185
|||||
223 ValProGluAlaIleAlaPheAspMetValGluIleAsnTyrTrpAspGl 239
1186 ACAGCATAAAACCTGTATGCTCAATGCCACATCAAAA.....TTCATGG 1229
|||||
239 nAspLeuTrpValCysMetLeuAlaSerGluGlnLysSerSerPheMetM 256
1230 AGTTCTACCAAGATGTAAGGACTGTGGCTTTCGGGTTCTATTCTGT 1279
|||||
256 euPheTyrArgAspValLysAspTrpLeuPheGlyPheTyrPheCys 272
1280 ATGCCCTTGGTGGCACTCGATCTTCTACACCCCTCATGACTTGTGAGAT 1329
|||||
273 LeuProLeuValCysThrGlyIlePheTyrThrLeuMetSerCysGluMe 289
```

```

1330 GTTGAACAGAAAGGAGTGCAGCTTGAGAAATGCCCTCAGTGAACATCTTA 1379
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
289 tLeuSerLysArgAsnGly...MetArgIleAlaLeuAsnAspHisMetL 305
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1380 AGCAGCGTCGAGAAAGTGGCAAAACAGATTTTCTGCTGGTGTGTAATTTT 1429
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
305 ysArgArgGluValAlaLysThrValPheCysLeuValValIlePhe 321
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1430 GCTCTTTCCTGGTTCCTCTTCACCTTAAGCCGTATATTGAAGAAACATGT 1479
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
322 AlaLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLysThrI 338
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1480 GTATAACCAAAATGGACAAAGACCGATGGAATTAATCTAGTTTCTTACTGC 1529
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
338 eTyAspGlnThrAspProAsnArgCysGluLeuLeuSerPheLeuLeu 355
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1530 TCATGGATTACATCGGTATTAACTTGGCAACCACTGAATTCATGTATAAAC 1579
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
355 alMetAspTyPheGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 371
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1580 CCCATAGCTCTGTATTGTCGAGCAAGAAATTTAAATTTGTTCCAGTC 1629
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
372 ProValAlaLeuTyPheValSerArgLysPheLysAsnCysPheGlnSe 388
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1630 ATGCTCTGCTGCTGCTGCTTACCAGTCCAAAGTCTGTAGTACCTCGGTCC 1679
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
388 tCysLeuCysTrpCysGlnArgProAlaLeuSerIleThrProThrA 405
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1680 CCATGAACGGAACAGCATCCAGTGGAGAACCAACCATGATCAAAACCAAC 1729
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
405 spGluLysGlySerValGlyLysTrpLysAlaAsnGlyGln...GluLeu 420
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1730 AACACAGACCGGAGCAGCCATAAG 1753
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
421 GlyLeuAspArgSerSerArg 428
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

seq_name: sp_vertebrate:Q9DGM2

```

seq_documentation_block:
ID Q9DGM2 PRELIMINARY; PRT; 426 AA.
AC Q9DGM2;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN RECEPTOR BL.
OS Brachydanio rerio (zebrafish) (zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Parichy D.M., Kelsh R.N., Mellgren E., Lopes S., Rawls J.,
RA Johnson S.L.;
RT "Mutational analysis of roles for endothelin receptor bl during neural
RT crest and pigment pattern development in the zebrafish, Danio rerio.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF275636; AAC00977.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECIP_F1_2; 1.
KW Receptor.
FT VARIANT 156 156 C -> *.
FT VARIANT 184 184 A -> E.
FT VARIANT 318 318 W -> *.
SQ SEQUENCE 426 AA; 48770 MW; F92C0B0B3046E4A3 CRC64;

```

alignment_scores:

Quality: 1260.50
Ratio: 3.718

Length: 440
Gaps: 10

```

Percent Similarity: 77.045 Percent Identity: 55.682
alignment_block:
US-09-931-157-1 x Q9DGM2
Align seg 1/1 to: Q9DGM2 from: 1 to: 426
485 ATGGAACCCCTTGGCTCAGGCGATCCCTTTGGCTGGCAGCTGGTTGGATG 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
7 MetGluThrArgCysValPheCysPheLeuPheLeu..... 18
535 TGTAATCAGTGATTAATCCTCGAGATACAGCACAAATCTAAGCAATCATG 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
19 .....LeuThrGluHisI 23
585 TG.....GATGATTTCACCACCTTTTCGT..... 607
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 leAlaValMetSerAlaGlnGlyLysAspPheAsnGlnSerArgLeuSer 39
608 .....GGCACAGAGCTCAGCTTCCTGGTTACCACTCATCA 642
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
40 MetGlyProLeuSerProThrGlnLysSerThrIleValIleGlyAsnG 56
643 ACCCACTAATTTGCTCCTACCAGCAATGGCTCA...ATGCACAACATATT 689
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
56 n...IleAsnGluSerMetProArgArgProLysValLeuProPrometC 72
690 GCCCACAGCAGCACTAAATTAATCTTTCAGCTTTCAAAATACATTAAACACTGTG 739
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
72 ysThrAspProThrGluIleArgAspThrPheLysTyIleAsnThrVal 88
740 ATATCTTGTAATTTTCATCGTGGGAATGCTGGGAATGCAACTCTGCT 789
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
89 ValSerCysLeuValPheValValGlyIleIleGlyAsnSerThrLeuLe 105
790 CAGGATCATTTACCAGAACAAATGTATCAGGAATGGCCCCAACCGCTGA 839
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
105 uArgIleIleTyLysAsnLysCysMetArgAsnGlyProAsnIleLeuI 122
840 TAGCCAGTCTTGGCCCTTGGAGACCTTATCTATGTGGTCATTGATCTCCCT 889
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
122 leAlaSerLeuAlaLeuGlyAspLeuLeuHisIleMetIleAspIlePro 138
890 ATCAATGATTTAAGCTGCTGGCTGGCGCTGCTTTTGATCACAATGA 939
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
139 IleAsnValTyLysLeuLeuAlaLysAspTrpPro..... 150
940 CTTTGGCGTATTTCTTTGCCAAGCTGTTCCTTTTCGAGAGCTCCTCGG 989
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 .PheGlyValGlyLeuCysLysLeuValProPheIleGlnLysThrServ 167
990 TGGGGATCACCGTCTCAACCTCTCGCTCTTAGTGTGACAGGTACAGA 1039
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 alGlyIleThrIleLeuSerLeuCysAlaLeuSerIleAspArgPheArg 183
1040 GCAGTTGCCCTCCCTGGATCGGTGTTTCAGGGAATGGGATTCCTTTGGTAAC 1089
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 AlaValSerSerTrpAsnArgIleLysGlyIleGlyValProLysTrpTh 200
1090 TGCCATTGAATTCCTCCATCTCGATCCCTGCTCTTATCTGCTGCCCATTC 1139
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
200 rAlaIleGluIleLeuIleTrpValLeuSerIleIleLeuAlaValP 217
1140 CTGAAGCGATTGGCTTCGTTCATGCTACCTTTGAATATAGGGGGAACAG 1189
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
217 roGluAlaIleAlaPheAspMetIleThrMetAspTyLysGlyGluGln 233
1190 CATAAAACCTGTATGCTCAATGCCACATCA.....AAATTCATGAGATT 1233
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 LeuArgIleCysLeuLeuHisProLysGlnArgIleLysPheMetGlnPh 250
1234 CTACCAAGATGTAAGAGCTGGGCTCTTCGGGGTCTTATTTCTGTATGC 1283
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```


seq_name: sp_mammal:Q28467

```

1542 TCGGTATTACTTGGCAACCATGAATTCATGTATAAACCCCATACCTCTG 1591
|||||
134 leGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIleAlaLeu 150
|||||
1592 TATTTTCGTGAGCAAGAAATTTAAATAATTCCTTCCAGTCATGCTCTGCTG 1641
|||||
151 TyrPheValSerIlysPheIlyAsnCysPheGlnSerCysLeuCysCys 167
|||||
1642 CTGCTGTATTACCACTCCAAAAGTCTGATG 1669
|||||
167 sCysCysTyrGlnSerIlySerLeuMet 176
|||||
seq_name: sp_human.Q16432

seq_documentation_block:
ID Q16432 PRELIMINARY; PRT; 187 AA.
AC Q16432;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN-A RECEPTOR DELTA 4..
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A..
RX MEDLINE=96190719; PubMed=8611157;
RA Miyamoto Y., Yoshimasa T., Arai H., Takaya K., Ogawa Y., Itoh H.,
RA Nakao K.;
RT "Alternative RNA splicing of the human endothelin-A receptor generates

```

RT Alternative RNA splicing of the human endocannabinoid receptor gene generates multiple transcripts.";
RL Biochem. J. 313:795-801(1996).
RL EMBL; S81542; AAB36326.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 1.
DR PROSITE; PS50262; G_PROTEIN_RECEPTOR_FL2; 1.
DR Receptor.
SQ SEQUENCE 187 AA; 20763 MW; 60E5564BDE607A36 CRC64;

```

alignment_scores:
  Quality: 958.00      Length: 182
  Ratio: 5.264        Gaps: 0
  Percent Similarity: 100.000  Percent Identity: 100.000

alignment_block:
  US-09-931-157-1 x Q16432  ..

Align seg 1/1 to: Q16432  from: 1 to: 187

485  ATGGAAACCTTTGGCTCAGGCGATCCTTTTGGCTGGCACTGGTTGGATG 534
|||||
1   MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyC 17
|||||

535  TGTATATCAGTGATATCCTGAGAGATACAGCACAAATCTAAGCAATCATG 584
|||||
17  svalllesrAspAnProGluArgTyrSerThrAsnLeuSerAsnHisv 34
|||||

585  TGGATGATTTCACCACTTTTTCGTGGGCACAGAGCTCAGCTTCCTGGTTACC 634
|||||

```

34 aAspAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
635 ACTCATCAACCACCAATATTTGGTCTACCCAGCAATGGCTCAATGCACAA 684
51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMetHisas 67
685 CTATTGCCACAGCAGACATAAATTAATTCAGCTTCAATACATACAA 734
67 ntyCysProGlnGlnThrLysIleThrSerAlaPheLysTyriLeAsnt 84
735 CTGTGATATCTGTACTATTTTCATCGTGGGAATGGTGGGAATCCAAC 784
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
785 CTGCTCAGGATCATTTACCAAGCAAAATGTATGAGGAATGCCCAACGC 834
101 LeuLeuArgIleIleTyTGlnAsnLysCysMetArgAsnGlyProAsnAl 117
835 GCTCATAGCAGCTTGGCCCTTGAGACCTTATCATGTGTCATGTATC 884
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyfValValIleAspL 134
885 TCCCTATCAATGATTTAAAGCTGCTGGCGCTGGCCCTTTTGATCAC 934
134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpPropheAspHis 150
935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCTCTTTTTCAGAGTC 984
151 AsnAspPheGlyValPheLeuLysLysLeuPhePropheLeuGlnLysSe 167
985 CTGCGTGGGATCACCGCTCTCAACCTCTGGCGCTCTAGTGTTCAG 1030
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAsp 182

seq_name: sp_mammal:Q9GMC5

seq_documentation_block:
ID Q9GMC5 PRELIMINARY; PRT; 135 AA.
AC Q9GMC5;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENDOTHELIN A RECEPTOR (FRAGMENT).
GN ETA.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RA Mershon J.L., McElvy S., Clark K.E.;
RT "Mechanism of Uterine Vascular Refractoriness to Endothelin-1 in Pregnant Sheep."
RL Submitted (AUG-2000) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF293847; AGO2228.1; .
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PRO0237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 135 135
SQ SEQUENCE 135 AA; 15844 MW; D1B9B9C402D73BA2 CRC64;

alignment_scores:
Quality: 709.00 Length: 135
Ratio: 5.291 Gaps: 0
Percent Similarity: 99.259 Percent Identity: 97.778

alignment_block:
US-09-931-157-1 x Q9GMC5

Align seg 1/1 to: Q9GMC5 from: 1 to: 135
1079 CCTTTGGTAACTGCCATTCGAAATTTGCTCTCCATCTGGATCCTGCTCTTAT 1128
1 ProLeuValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIl 17
1129 CCTGGCCATTCTCTGAAGCGATTGGCTTCGTCATGGTACCCCTTTGAATATA 1178
17 eLeuAlaIleProGluAlaIleGlyPheValMetValPropheGluTyrl 34
1179 GGGGTGAACAGCATAAACCTCTGCTCAATGCCACATCAAAATTCATG 1228
::: 1179 GGGGTGAACAGCATAAACCTCTGCTCAATGCCACATCAAAATTCATG 1228
34 ysGlyAlaGlnHisArgThrCysMetLeuAsnAlaThrSerLysPheMet 50
1229 GAGTCTTACCAAGATGTAAGGACACGGTGGCTCTTCGGGTTCTATTTCTG 1278
51 GluPheTyfGlnAspValLysAspTrpTrpLeuPheGlyPheTyfPheCy 67
1279 TATGCCCTTGGTGTGCACTGGCCTCTTCTACACCCCTCATGACTTGTGACA 1328
67 sMetProLeuValCysThrAlaIlePheTyfThrLeuMetThrCysGluM 84
1329 TGTGAACAGAGGAATGGCAGCTTGAGAAATTCGCTCAGTGAACATCTT 1378
84 etLeuAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeu 100
1379 AAGCAGCGTCGAGAAGTGGCAAAACAGTTTCTGCTGGTGTGTAATTT 1428
101 LysGlnArgArgGluValAlaLysThrValPheCysLeuValValIlePh 117
1429 TGCTCTTTGCTGGTTCCTCTTCACTTAAGCGGTATATTGAAGAAACATG 1478
117 eAlaLeuCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrv 134
1479 TGTAT 1483
134 altYr 135

OM of: US-09-931-157-1 to: SwissProt_40.* out_format : pfs

Date: Oct 1, 2002 1:21 AM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+2np.model -DEV=xlp
-O=/cgn2_1/uspt0_pool/US09931157/tunat_30092002_110421_11329/app_query.fasta_1.8657
-DB=SwissProt_40 -OFMT=fastan -SUFFIX=rsp -GAPOP=12.000
-CAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09931157 -CGN1_1_127 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-931-157-1

Query length: 4105

Database: SwissProt_40.*

Database sequences: 105224

Database length: 38719550

Search time (sec): 126.310000

score_list:

Sequence	Strd Orig	Zscore	Escore	Len	Documentation
SwissProt_40:ETIR_HUMAN	2282.00	2936.67	1.6e-156	427	P25101 homo sapiens (human).
SwissProt_40:ETIR_PIG	2187.00	2814.01	1.1e-149	427	Q29010 sus scrofa (pig). endot
SwissProt_40:ETIR_BOVIN	2155.00	2772.69	2.2e-147	427	P21450 bos taurus (bovine).
SwissProt_40:ETIR_RAT	2108.50	2712.68	4.9e-144	426	P26684 rattus norvegicus (rat).
SwissProt_40:ETIR_MOUSE	1391.00	1790.43	1.8e-92	270	Q61614 mus musculus (mouse).
SwissProt_40:ETIR_PIG	1254.50	1609.67	1.3e-82	443	P35463 sus scrofa (pig). endot
SwissProt_40:ETIR_COTJA	1246.50	1601.97	4.7e-81	347	Q90328 coturnix coturnix (chick).
SwissProt_40:ETIR_HUMAN	1240.00	1590.97	1.4e-81	442	P24530 homo sapiens (human).
SwissProt_40:ETIR_RAT	1240.00	1590.97	1.4e-81	442	P21451 rattus norvegicus (rat).
SwissProt_40:ETIR_MOUSE	1236.50	1586.45	2.6e-81	442	P48302 mus musculus (mouse).
SwissProt_40:ETIR_CANFA	1228.00	1575.82	1.0e-80	426	P56497 canis familiaris (dog).
SwissProt_40:ETIR_HORSE	1226.00	1572.88	1.5e-80	443	Q62709 equus caballus (horse).
SwissProt_40:ETIR_BOVIN	1221.50	1567.11	3.1e-80	441	P28088 bos taurus (bovine).
SwissProt_40:ETIR_XENLA	1014.50	1299.78	2.4e-65	444	P32940 xenopus laevis (afric).
SwissProt_40:GRPR_MOUSE	466.00	592.90	6.5e-26	384	P21729 mus musculus (mouse).
SwissProt_40:GRPR_RAT	463.00	589.03	1.1e-25	384	P52500 rattus norvegicus (rat).
SwissProt_40:GRPR_HUMAN	455.50	579.35	3.7e-25	384	P30550 homo sapiens (human).
SwissProt_40:NMBR_MOUSE	454.50	577.91	4.4e-25	390	P24053 rattus norvegicus (rat).
SwissProt_40:NMBR_HUMAN	450.50	572.75	8.5e-25	390	O54799 mus musculus (mouse).
SwissProt_40:NMBR_MOUSE	434.50	552.09	1.2e-23	390	P28336 homo sapiens (human).
SwissProt_40:BR33_MOUSE	427.50	542.84	3.8e-23	399	O54798 mus musculus (mouse).
SwissProt_40:BR33_CAVPO	422.00	535.74	9.5e-23	399	P35371 cavia porcellus (guin).
SwissProt_40:BR33_SHEEP	419.50	532.51	1.4e-22	399	O97967 ovis aries (sheep).
SwissProt_40:BR34_HOMOR	413.50	520.85	4.2e-22	613	O15354 homina orientalis (c).
SwissProt_40:BR34_HOMOR	411.00	521.70	5.9e-22	392	P47751 homina orientalis (c).
SwissProt_40:BR33_HUMAN	409.00	518.96	8.2e-22	399	P32247 homo sapiens (human).
SwissProt_40:EBP2_HUMAN	367.00	463.03	8.9e-19	481	O60883 homo sapiens (human).
SwissProt_40:OPRM_BOVIN	335.00	423.37	1.7e-16	401	P79350 bos taurus (bovine).
SwissProt_40:OPRX_CAVPO	329.50	417.00	4.2e-16	370	P47748 cavia porcellus (guin).
SwissProt_40:ETBR_MACFA	328.00	427.08	4.3e-16	99	Q28468 macaca fascicularis (c).
SwissProt_40:OPRM_MOUSE	322.50	407.30	1.4e-15	398	P42866 mus musculus (mouse).
SwissProt_40:OPRM_RAT	322.50	407.30	1.4e-15	398	P33535 rattus norvegicus (rat).
SwissProt_40:OPRX_PIG	317.50	401.50	3.1e-15	370	P79292 sus scrofa (pig). nocid
SwissProt_40:OPRM_MACMU	316.50	399.50	3.7e-15	400	Q9myw9 macaca mulatta (rhesu).
SwissProt_40:OPRM_HUMAN	315.50	398.21	4.3e-15	400	P35372 homo sapiens (human).
SwissProt_40:OPRX_MOUSE	315.00	398.35	4.7e-15	367	P35377 mus musculus (mouse).
SwissProt_40:OPRX_HUMAN	313.50	396.34	6.0e-15	370	P41146 homo sapiens (human).
SwissProt_40:OPRX_MOUSE	313.00	395.45	6.5e-15	380	P33534 mus musculus (mouse).
SwissProt_40:OPRX_RAT	312.00	394.48	7.7e-15	367	P35370 rattus norvegicus (rat).
SwissProt_40:OPRM_PIG	312.00	393.67	7.8e-15	401	Q95247 sus scrofa (pig). mu-ty

```
SwissProt_40:OPRK_RAT + 310.00 391.58 1.1e-14 380 ! P34975 rattus norvegicus
SwissProt_40:GHSR_HUMAN + 305.00 385.12 2.5e-14 380 ! P41145 homo sapiens (h
SwissProt_40:GHSR_PIG + 304.50 384.82 2.6e-14 366 ! Q95254 sus scrofa (pig).
SwissProt_40:GHSR_HUMAN + 302.50 382.24 3.7e-14 366 ! Q92847 homo sapiens (h
SwissProt_40:OPRK_CAVPO + 302.00 381.25 4.0e-14 380 ! P41144 cavia porcellus

seq_name: SwissProt_40:ETIR_HUMAN

seq_documentation_block:
ID ETIR_HUMAN STANDARD; PRT; 427 AA.
AC P25101;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endothelin-1 receptor precursor (ET-A).
GN EDNR A OR ETRA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92062153; PubMed=1719979;
RA Adachi M., Yang Y.Y., Furuichi Y., Miyamoto C.;
RT "Cloning and characterization of cDNA encoding human A-type
endothelin receptor.";
RL Biochem. Biophys. Res. Commun. 180:1265-1272(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=92068188; PubMed=1659806;
RA Cyr C., Huebner K., Druck T., Kris R.;
RT "Cloning and chromosomal localization of a human endothelin ETA
receptor.";
RL Biochem. Biophys. Res. Commun. 181:184-190(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91348221; PubMed=1652463;
RA Hosoda K., Nakao K., Arai H., Suga S., Ogawa Y., Mukoyama M.,
Shirakami G., Saito Y., Nakanishi S., Imura H.;
RT "Cloning and expression of human endothelin-1 receptor cDNA.";
RL FEBS Lett. 287:23-26(1991).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93180293; PubMed=1291713;
RA Arai H., Nakao K., Hosoda K., Ogawa Y., Nakagawa O., Komatsu Y.,
Imura H.;
RT "Molecular cloning of human endothelin receptors and their expression
in vascular endothelial cells and smooth muscle cells.";
RL Jpn. Circ. J. 56:1303-1307(1992).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=93179382; PubMed=8440682;
RA Elshourbagy N.A., Korman D.R., Wu H.L., Sylvester D.R., Lee J.A.,
Nuthalaganti P., Bergsma D.J., Kumar C.S., Nambi P.;
RT "Molecular characterization and regulation of the human endothelin
receptors.";
RL J. Biol. Chem. 268:3873-3879(1993).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=92406798; PubMed=1326535;
RA Hosoda K., Nakao K., Tamura N., Arai H., Ogawa Y., Suga S.,
Nakanishi S., Imura H.;
RT "Organization, structure, chromosomal assignment, and expression of
the gene encoding the human endothelin-A receptor.";
RL J. Biol. Chem. 267:18797-18804(1992).
RN [7]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=93035452; PubMed=1415318;
RA Hayzer D.J., Rose P.M., Lynch J.S., Webb M.L., Kienzle B.K., Liu E.C.,
Bogosian E.A., Brinson E., Runge M.S.;
```


RT "Cloning and expression of a human endothelin receptor: subtype A.";
 RL Am. J. Med. Sci. 304:231-238(1992).
 RP [8]
 RN SEQUENCE FROM N.A.
 RC TISSUE=Lung;
 RX MEDLINE=96190719; PubMed=8611157;
 RA Miyamoto Y., Yoshimasa T., Arai H., Takaya K., Ogawa Y., Itoh H.,
 Nakao K.;
 RT "Alternative RNA splicing of the human endothelin-A receptor
 generates multiple transcripts.";
 RL Biochem. J. 313:795-801(1996).
 RP [9]
 RN SEQUENCE OF 1-140 FROM N.A.
 RX MEDLINE=93151811; PubMed=8427579;
 RA Yang H., Tabuchi H., Furuichi Y., Miyamoto C.;
 RT "Molecular characterization of the 5'-flanking region of human
 genomic ETA gene.";
 RL Biochem. Biophys. Res. Commun. 190:332-339(1993).
 CC -1- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
 ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
 CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
 CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D90348; BAA14359.1; -;
 DR EMBL; L06622; AAA58447.1; -;
 DR EMBL; X61950; CAA43953.1; -;
 DR EMBL; S67127; AAB20407.1; -;
 DR EMBL; S63938; AAB20278.1; -;
 DR EMBL; S57498; CAB31573.1; ALT_SEQ.
 DR EMBL; D11151; BAA01920.1; -;
 DR EMBL; D11145; BAA01920.1; JOINED.
 DR EMBL; D11146; BAA01920.1; JOINED.
 DR EMBL; D11147; BAA01920.1; JOINED.
 DR EMBL; D11148; BAA01920.1; JOINED.
 DR EMBL; D11149; BAA01920.1; JOINED.
 DR EMBL; D11150; BAA01920.1; JOINED.
 DR EMBL; S45956; AAB23644.1; -;
 DR EMBL; S81539; AAB36325.1; -;
 DR EMBL; S55772; AAB25212.1; -;
 DR PIR; JS0607; JS0607.
 DR PIR; JT0972; JT0972.
 DR PIR; S16980; S16980.
 DR PIR; A44158; A44158.
 DR PIR; A45481; A45481.
 DR GCRDB; GCR_0403; -;
 DR GCRDB; GCR_0473; -;
 DR GCRDB; GCR_0491; -;
 DR GCRDB; GCR_0512; -;
 DR GCRDB; GCR_0602; -;
 DR GCRDB; GCR_0727; -;
 DR GCRDB; GCR_0737; -;
 DR GCRDB; GCR_1922; -;
 DR MIM; 131243; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
 DR PROSITE; PS02662; G_PROTEIN_RECEP_FL_2; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 KW SIGNAL
 FT CHAIN 1 17
 FT CHAIN 18 427
 FT CHAIN 18 80
 FT DOMAIN 81 102
 FT TRANSMEM 1 (POTENTIAL).

FT DOMAIN 103 112 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 113 132 2 (POTENTIAL).
 FT DOMAIN 133 159 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 160 181 3 (POTENTIAL).
 FT DOMAIN 182 205 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 206 229 4 (POTENTIAL).
 FT DOMAIN 230 256 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 257 278 5 (POTENTIAL).
 FT DOMAIN 279 306 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 307 328 6 (POTENTIAL).
 FT DOMAIN 329 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 372 7 (POTENTIAL).
 FT DOMAIN 373 427 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 158 239 BY SIMILARITY.
 FT CONFLICT 110 110 C -> Y (IN REF. 7).
 FT CONFLICT 207 207 V -> A (IN REF. 5).
 FT CONFLICT 234 234 E -> G (IN REF. 5).
 FT CONFLICT 280 280 C -> G (IN REF. 5).
 SQ SEQUENCE 427 AA; 48722 MW; 207272D4A231404F CRC64;

alignment_scores:
 Quality: 2282.00 Length: 427
 Ratio: 5.344 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-931-157-1 x ETIR_HUMAN ..
 Align seg 1/1 to: ETIR_HUMAN from: 1 to: 427

485 ATGGAACCCCTTTGCTCAGGCGATCCTTTTGGCTGGCACTGGTTGGATG 534
 1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyCy 17
 535 TGTAAATCAGTGATAATCTCTGAGAGATACACACAAATCTAAGCAATCATG 584
 17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisV 34
 585 TGGATGATTTCCACACACTTTTCGTGGGCACAGCTCAGCTTCCTGGTTACC 634
 34 aAlAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
 635 ACTCATCAACCCCACTAATTTGGTCTACCCAGCAATGGCTCAATGCACAA 684
 51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMetHisAs 67
 685 CTATTGCCACACAGACTAAATTTACTTCAAGCTTTCAATATACATTACA 734
 67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
 735 CTGTGATATCTGTACTATTTTCATCGTGGGAATGGTGGGAATGCAACT 784
 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
 785 CTGCTCAGGATCATTTACCAAGAACAAATGTATAGGAATGGCCCAACGC 834
 101 LeuLeuArgIleIleIleYrGlnAsnLysCysMetargAsnGlyProAsnAl 117
 835 GCTGATAGCCAGCTCTGCCCTTGGAGACCTTATCTATGTGTGTCATTGATC 884
 117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
 885 TCCCTATCAATGATTTAAGCTGCTGGCGCTGGCGCTGGCCCTTTTGATCAC 934
 134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHis 150
 935 AATGACATTTGGCGTATTTCTTTCGAAGCTGTTCCCTTTTTCGACGAAGTC 984
 151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSe 167

```
985 CTCGGTGGGATACCGCTCAACCTCTCGGCTCTTAGTGTGACAGGT 1034
|||||
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
|||||
1035 ACAGAGCAGTGGCTCTGGAGTGTCTCAGGGAATGGCATTCCTTTG 1084
|||||
184 yArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
|||||
1085 GTAACTGCCATTGAATGTCTCCATCTGGATCTGCTCTTATCTGATGCC 1134
|||||
201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
|||||
1135 CATTCCTGAACGATGTGCTTCGTCATCGGTACCTTTCGAATATAGGGGT 1184
|||||
217 aileProGluAlaIleGlyPheValMetValProPheGluTrpArgGly 234
|||||
1185 AACAGCAATAACCTGTATGCTCAATGCCACATCAAAATTCATGAGTTC 1234
|||||
234 luGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
|||||
1235 TACCAAGATGTAAAGGACTGGTGGCTCTTCGGGTCTTATTTCTGATGCC 1284
|||||
251 TyrGlnaspValLysAspTrpLeuPheGlyPheThrPheCysMetPr 267
|||||
1285 CTTGGTGGCACTCGGATCTTCTACACCTCATGACTGTGTGATGTGA 1334
|||||
267 oLeuValCysThrAlaIlePheThrLeuMetThrCysGluMetLeuA 284
|||||
1335 ACAGAAGCAATGGCAGCTGTGAGATTCGCCCTCAGTGAACATCTTAAGCAG 1384
|||||
284 snArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
|||||
1385 CGTCGAGAACTGGCAAAACAGTTCTTCGCTGTGTTGTAATTTTGTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
|||||
1435 TTGCTGGTTCCTCTTCACTTAAGCCGTATATTGAAGAAACTGTGTATA 1484
|||||
317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyra 334
|||||
1485 ACGAAATGGCAAGAACCGATGTGAATTAATTTCTTACTGCTCATG 1534
|||||
334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350
|||||
1535 GATTACATCGGTATTAATTCGGCAACCATGAATTCATGTATATAACCCCAT 1584
|||||
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
|||||
1585 AGCTCTGTATTTTGTGACCAAGAATTTAAAAATGTTTCCAGTCATGCC 1634
|||||
367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
|||||
1635 TCTGCTGCTGTGTATACCAAGTCCAAAGCTCTGATGACCTCGTCCCATG 1684
|||||
384 euCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValProMet 400
|||||
1685 AACGGAACAACATCCAGTGGAGAACACCATCAAAACCAACCAACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnAsnHisAsnTh 417
|||||
1735 AGACCGGAGCAGCATATAGGACACATGAAC 1765
|||||
417 rAspArgSerSerHisLysAspSerMetAsn 427
|||||
seq_name: SwissProt_40:ETIR_PIG
seq_documentation_block:
ID ETIR_PIG STANDARD; PRT; 427 AA.
AC Q29010;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Endothelin-1 receptor precursor (ET-A).
```

```
GN EDNR.
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=96031054; PubMed=8564193;
RA Nishimura J., Aoki H., Chen X., Shikasho T., Kobayashi S.,
RA Kanade H.;
RT "Evidence for the presence of endothelin ETA receptors in endothelial
RT cells in situ on the aortic side of porcine aortic valve.";
RL Br. J. Pharmacol. 115:1369-1376(1995).
CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR ENBL; S80652; AAB36014.1;
DR GCRD5; GCR_1954;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEPTOR_FL1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEPTOR_FL2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1
FT CHAIN 18 427 ENDOTHELIN-1 RECEPTOR.
FT DOMAIN 81 102 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 103 112 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 113 132 2 (POTENTIAL).
FT TRANSMEM 133 159 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 160 181 3 (POTENTIAL).
FT TRANSMEM 182 205 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 206 229 4 (POTENTIAL).
FT DOMAIN 230 256 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 257 278 5 (POTENTIAL).
FT DOMAIN 279 306 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 307 328 6 (POTENTIAL).
FT DOMAIN 329 347 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 348 372 7 (POTENTIAL).
FT DOMAIN 373 427 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
FT DISULFID 158 239 BY SIMILARITY.
SQ SEQUENCE 427 AA; 48694 MW; F09366051552D06D CRC64;
alignment_scores:
Quality: 2187.00 Length: 427
Ratio: 5.134 Gaps: 0
Percent Similarity: 99.766 Percent Identity: 94.848
alignment_block:
US-09-931-157-1 x ETIR_PIG
Align seg 1/1 to: ETIR_PIG from: 1 to: 427
485 ATGGAACCCCTTTCCTCAGGCATCCCTTTGGCTGGCAGCTGGTGGATG 534
|||||
1 MetGluThrPheCysPheArgValSerPheTrpValAlaLeuLeuGlyC 17
```

```
535 TGAATCACTGATAATCTCGAGAGATACAGACAAATCTAAGCAATCATG 584
|||||
17 sValileSerAspAsnProGluSerHisSerThrAsnLeuSerThrHisv 34
|||||
585 TGGATGATTTCACCACTTTTCGTGGCAGAGACTCAGCTCCCTGGTTACC 634
|||||
34 aAspAspPheThrThrPheArggLyThrGluPheSerLeuValValThr 50
|||||
635 ACTCATCAACCCACTAATTTGGTCTCACCAGCAATGGCTCAATGCACAA 684
|||||
51 ThrHisArgProThrAsnLeuAlaLeuProSerAsnGlySerMetHisas 67
|||||
685 CTATTGCCACAGCAGACTAAATTTACTTCAGCTTTCAATCAATTAACA 734
|||||
67 nYrCysProGlnGlnThrLysIleThrSerAlaPheLysTyIleAsnt 84
|||||
735 CNGTATATCTTGTACTATTTTCATCGTGGGATGGTGGGAATGCAACT 784
|||||
84 hrValileSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
|||||
785 CTGCTCAGGATCATTTACCAGAACAAATGATGAGGAATGGCCCAACGC 834
|||||
101 LeuLeuArgIleIleIleTyIleGlnAsnLysCysMetArgAsnGlyProAsnAl 117
|||||
835 GCTGATAGCAGCTTGTCCCTTGGAGACCTTATCTATGTGGTCATTCATC 884
|||||
117 aleuileAlaSerLeuAlaLeuGlyAspLeuIleTyIleValIleAspL 134
|||||
885 TCCCTATCAATGTATTAAGTCTGGCTGGGCGCTGTCCTTTGATGCAC 934
|||||
134 euProileAsnValPheLysLeuLeuAlaGlyArgTrpProPheGluAsn 150
|||||
935 ATGACTTGGCGTATTTCTTCCAGCTGTCCTCCCTTTTTCGAGAAGTC 984
|||||
151 HisAspPheGlyValPheLeuLysLysLeuPheProPheLeuGlnLys 167
|||||
985 CTCGCTGGGATCACCGCTCTCAACCTCTCGCTCTTAGTTGTGACAGGT 1034
|||||
167 rSerValGlyIleThrValLeuAsnLeuLysAlaLeuSerValAspArgT 184
|||||
1035 ACAGACGAGTTCCTCTCGAGTCGTGTCAGGGAATGGGATTCCTTTG 1084
|||||
184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
|||||
1085 GTAACCTGCCATGAATGTCTCCATCTGGATCTGTCCTTTATCTCTGGC 1134
|||||
201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
|||||
1135 CATTCCTGAAGCGATTGGCTTCGTCTATGTTACCTTTTGAATATAGGGTG 1184
|||||
217 alileProGluAlaIleGlyPheValMetValProPheGluTyIleGlyG 234
|||||
1185 ACAGCATAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
|||||
234 luGluHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
|||||
1235 TACCAAGATGAAGCAGCTGGTGGCTCTTCGGCTCTATTCTGTATGCC 1284
|||||
251 TyIleAspValLysAspTrpTrpLeuPheGlyPheTyIlePheCysMetPr 267
|||||
1285 CTGTGTGTGCACTCGCATCTTCACACCTCATGACTGTGTGAGATGTTGA 1334
|||||
267 OleuValCysThrAlaIlePheTyIleThrLeuMetThrCysGluMetLeuA 284
|||||
1335 ACAGAGGAATGCCAGCTTGAGAAATGCCCTCAGTGAACATCTTAAGCAG 1384
|||||
284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
|||||
1385 CGTCGAGAGTGCACAAACAGTTTCTGCTGCTGCTGCTGCTGCTGCTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
|||||
```

```
1435 TTGCTGGTTCCTCTTCACTTAAGCCGTATATTCAAGAAACTGTGTATA 1484
|||||
317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyra 334
|||||
1485 ACCAAATGGACAAGACCGATGTGAATTACTTAGTTCTTACTGCTCATG 1534
|||||
334 spGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350
|||||
1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAACCCCAT 1584
|||||
351 AspTyIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
|||||
1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAATTTGTTTCCAGTCATGCC 1634
|||||
367 eAlaLeuTyIlePheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
|||||
1635 TCTGCTGCTGCTGTACCACTCCAAAAGTCTGATGACCTCGTCCCGCATG 1684
|||||
384 euCysCysCysTyIleGlnSerLysSerLeuMetThrSerValProMet 400
|||||
1685 AACCGAACAAGCATCCAGTGGAGAAGACCAAGATCAAAACACCAACACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnHisGluGlnAsnHisAsnTh 417
|||||
1735 AGACCGAGCAGCATTAAGACAGCATGAAC 1765
|||||
417 rGluArgSerSerHisLysAspSerIleAsn 427
|||||
seq_name: SwissProt_40:ETIR_BOVIN

seq_documentation_block:
ID ETIR_BOVIN STANDARD; PRT; 427 AA.
AC P21450;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Endothelin-1 receptor precursor (ET-A).
ED EDNA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_Taxid=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=91080923; PubMed=2175396;
RA Arai H., Hori S., Aramori I., Ohkubo H., Nakanishi S.;
RT "Cloning and expression of a cDNA encoding an endothelin receptor.";
RL Nature 348:730-732(1990).
CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X57765; CAA40917.1; -
DR PIR; S13424; S13424.
DR GCRB; GCR_0005; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PR00237; GPCR_RHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
```

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 17 POTENTIAL.
 FT CHAIN 18 427 ENDOTHELIN-1 RECEPTOR.
 FT DOMAIN 18 80 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 81 102 1 (POTENTIAL).
 FT DOMAIN 103 112 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 113 132 2 (POTENTIAL).
 FT DOMAIN 133 159 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 160 181 3 (POTENTIAL).
 FT DOMAIN 182 205 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 206 229 4 (POTENTIAL).
 FT DOMAIN 230 256 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 257 278 5 (POTENTIAL).
 FT DOMAIN 279 306 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 307 328 6 (POTENTIAL).
 FT DOMAIN 329 347 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 348 372 7 (POTENTIAL).
 FT DOMAIN 373 427 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 62 62 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 158 239 BY SIMILARITY.
 SQ SEQUENCE 427 AA; 48515 MW; 73CF4404CD9AC14 CRC64;

alignment_scores:

Quality: 2155.00 Length: 427
 Ratio: 5.143 Gaps: 0
 Percent Similarity: 98.126 Percent Identity: 94.614

alignment_block:

US-09-931-157-1 x ETIR_BOVIN

Align seg 1/1 to: ETIR_BOVIN from: 1 to: 427

485 ATGGAACCCCTTGCCTCAGGCGATCCTTTTGGCTGGCAGCTGGTGGATG 534
 |||||..... |||||.....
 1 MetGluThrPheTrpLeuArgLeuSerPheTrpValAlaLeuValGlyGI 17
 535 TGTATCAGTATCTCTGAGATACAGACACAAATCTAGCAATCATG 584
 |||||..... |||||.....
 17 yValIleSerAspAsnProGluSerTyrSerThrAsnLeuSerIleHisV 34
 585 TGGATGATTCCACCACTTTTCGTGGCAGAGCTCAGCTTCTGGTTACC 634
 |||||..... |||||.....
 34 alaSPSerValAlaThrPheHisGlyThrGluLeuSerPheValValThr 50
 635 ACTCATCAACCCACTAATTTGGTCCTACCCAGCAATGGCTCAATGCCAA 684
 |||||..... |||||.....
 51 ThrHisGlnProThrAsnLeuAlaLeuProSerAsnGlySerMetHisAs 67
 685 CTATTGCCAGCAGACTAAATTTACTTTCAGCTTTTCAATACATTACA 734
 |||||..... |||||.....
 67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
 735 CTGTGATATCTGTACTATTTTCATCTCGTGGGAATGGGGAATGCAACT 784
 |||||..... |||||.....
 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
 785 CTGCTCAGGATCATTTTACCAGAACAAATGTATGAGGAATGCCCCCAACGC 834
 |||||..... |||||.....
 101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
 835 GCTGATACCCAGCTTTCGCCCTTGGAGACCTTATCTATGTGCTCATTCATC 884
 |||||..... |||||.....
 117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValIleAspL 134
 885 TCCCTATCATCTATTAAAGCTGTGCTGGCGCTGGCGCTTGTGATCAC 934
 |||||..... |||||.....
 134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheGluGln 150
 935 AATGACTTTGGCGTATTTCTTTTGAAGCTGTTCCCTTTTGGCAGAAGTC 984
 |||||..... |||||.....
 151 AsnAspPheGlyValPheLeuLysLeuPheProPheLeuGlnLysSe 167

985 CTCGGTGGGATCACCCGCTCCTCAACCTCTGGCTCTTAGTGTGTGACAGT 1034
 |||||..... |||||.....
 167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
 1035 ACAGAGCAGTTGCTCCTCCTGGAGTCTGTCTAGGGAATGGGATTCCTTTG 1084
 |||||..... |||||.....
 184 yArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
 1085 GTAACCTGCCAATGAAATGTCTCCATCTGGATCTCTCTCTTATCTCTGC 1134
 |||||..... |||||.....
 201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
 1135 CATTCTCTGAACCGATTGGCTTCGTGATACCGCTTTCGAAATATAGGGGTG 1184
 |||||..... |||||.....
 217 aIleProGluAlaIleGlyPheValMetValProPheGluTyrLysGlyA 234
 1185 AACAGCAATAAACCTGTATGCTCATGCCATCAATCAAAATTCATGAGTTC 1234
 |||||..... |||||.....
 234 laGlnHisArgThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
 1235 TACCAAGATGTAAGGACTGGTGGCTCTTCGGGTTCTATTCTGTATGCC 1284
 |||||..... |||||.....
 251 TyrGlnAspValLysAspTrpIlePheGlyPheTyrPheCysMetPr 267
 1285 CTTGGTGTGCACTGGGATCTTCTACACCTCATGACTTGTGAGATGTGA 1334
 |||||..... |||||.....
 267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
 1335 ACAGAAGGAATGGCAGCTTGAGATTGCCCTCAGTGAACATCTTAAGCAG 1384
 |||||..... |||||.....
 284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
 1385 CGTCGAGAAGTGGCAAAACAGTTTCTCTGCTGGTGTGAATTTTGTCTCT 1434
 |||||..... |||||.....
 301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
 1435 TTGCTGTTCCCTCTTCACTTAAGCCGTATATTGAGAAACTGTGTATA 1484
 |||||..... |||||.....
 317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
 1485 ACGAAATGGCAAGAACCGATGTGAATTAATTAATTTCTTCTTCTGCTCATG 1534
 |||||..... |||||.....
 334 spGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350
 1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAAACCCTAT 1584
 |||||..... |||||.....
 351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIle 367
 1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTTCCAGTCATGCC 1634
 |||||..... |||||.....
 367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
 1635 TCTGCTGCTGCTTACCAGTCCAAAGCTCTGATGACCTCGTCCGCCATG 1684
 |||||..... |||||.....
 384 euCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValProMet 400
 1685 AACGGAACAAGCATCCAGTGGGAAGAACCATCAAAACCAACACCAACAC 1734
 |||||..... |||||.....
 401 AsnGlyThrSerIleGlnTrpLysAsnHisGluGlnAsnAsnHisAsnTh 417
 1735 AGACCGGAGCAGCCATAGGACACATGAAC 1765
 |||||..... |||||.....
 417 rGluArgSerSerHisLysAspSerIleAsn 427

seq_name: SwissProt_40:ETIR_RAT

seq_documentation_block:

ID ETIR_RAT STANDARD: PRT: 426 AA.
 AC P26684;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-NOV-1995 (Rel. 32, Last annotation update)


```

|||||
251 TyrClnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPr 267
1285 TTGGTGTGCAGTCGATCTTACACCCCTCATGCTTGAGATGTTGA 1334
|||||
267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
1335 ACAGAAGGAATGCGAGCTTCAGAAATTCGCCCTCAGTGAACATCTTAAGCAG 1384
|||||
284 snArgArgAnglySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
1385 CGTCGAGAAGTGGCAAAACAGATTTCTGCTTGGTTGTAATTTTGTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 317
1435 TTGCTGTTCCCTCTTACCTTAAGCCGTATATTGAAGAAACCTGTGTATA 1484
|||||
317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 ACAGAAATGGACAAGACCGATGTGAATTACTTACTTCTTACTGCTCATG 1534
|||||
334 spGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
1535 GATTACATCGGTATTAATCTGGCAACCATGAATTCATGATATAAACCCT 1584
|||||
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
1585 AGCTCTCTATTGCTGACGAGAAATTTAAATTTGTTCCAGTCATGCC 1634
|||||
367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TCCTGCTCTGCTTACCATGTCACCAAGTCTGTGACCTCGGTCCTCCATG 1684
|||||
384 euCysCysCysCysHisGlnSerLysSerLeuMetThrSerValProMet 400
1685 AACGGAACAGCATCCAGTGGAGAACCAGCATCAAAACACCAACACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnGlnGln...AsnHisAsnTh 416
1735 AGACCGAGCAGCATTAAGACACGATGAAC 1765
|||||
416 rGluArgSerSerHisLysAspSerMetAsn 426

```

seq_name: SwissProt_40-ETIR_MOUSE

```

seq_documentation_block:
ID ETIR_MOUSE STANDARD; PRT; 270 AA.
AC O61614; O54993;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endothelin-1 receptor (ET-A) (ET-AR) (Fragment).
GN EDNRA OR GPCR10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Lung;
RA Schoenfeld J.R., Lowe D.G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 100-216 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94116980; PubMed=8288218;
RA Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,
RA Copeland N.G., Jenkins N.A.;
RT "Identification, chromosomal location, and genome organization of
RL mammalian G-protein-coupled receptors.";
RL Genomics 18:175-184(1993).
CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-

```

```

CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF039892; AAB96671.1;
CC EMBL: L20340; AAA16845.1;
CC GCRDB; GCR_0923;
CC MGD; MGI:105923; Ednra.
CC MGD; MGI:99570; Gpcr10.
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1.1.
CC PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
CC PROSITE; PS00462; G_PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT NON_TER 1
FT TRANSMEM 1
FT TRANSMEM <1 4 1 (POTENTIAL).
FT DOMAIN 5 14 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 15 34 2 (POTENTIAL).
FT DOMAIN 35 61 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 62 83 3 (POTENTIAL).
FT DOMAIN 84 107 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 108 131 4 (POTENTIAL).
FT DOMAIN 132 158 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 159 180 5 (POTENTIAL).
FT DOMAIN 181 208 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 209 230 6 (POTENTIAL).
FT DOMAIN 231 249 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 250 >270 7 (POTENTIAL).
FT DISULFID 60 141 BY SIMILARITY.
FT CARBOHYD 144 144 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT NON_TER 270
SQ SEQUENCE 270 AA; 30972 MW; 34C76B74B70BC99E CRC64;

```

alignment_scores:

Quality: 1391.00 Length: 270
Ratio: 5.190 Gaps: 0
Percent Similarity: 99.259 Percent Identity: 97.778

alignment_block:

US-09-931-157-1 x ETIR_MOUSE

Align seg 1/1 to: ETIR_MOUSE from: 1 to: 270

```

779 GCAACTCTGCTCAGGATCATTTACGAGCAAAATGTATGAGGAATGGCCC 828
|||||
1 AlaThrLeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyPr 17
829 CAACGCGCTGATACCGACTCTTGGCCCTTGGAGACCTTATCTATGTGTGCA 878
|||||
17 oAsnAlaLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValI 34
879 TTGATCTCCCTATCAATGTATTAAAGCTGTGCTGGCTGGCGCTGGCCTTTT 928
|||||
34 leAspLeuProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPhe 50
929 GATCACAAATGACTTTGGCGTATTTCTTGTGCAAGCTGTTCCCTTTTTCGA 978
|||||
51 AspHisAsnAspPheGlyValPheLeuLeuLysLysLeuPheProPheLeuGl 67
979 GAAGTCCTCGGTGGGATCACCGTCCTCAACCTCTCGCCTCTTAGTGTG 1028
|||||
67 nLysSerSerValGlyIleThrValLeuAsnLeuLysCysAlaLeuSerValA 84

```

CC	MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE	
CC	A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.	
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.	
CC	-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.	
DR	GRdbb; GCR_09337; -	
DR	InterPro; IPR000276; GPCR_Rhodpsn.	
DR	Pfam; PF00001; 7tm_1; 1.	
DR	PRINTS; PR00237; GPCRHHODPSN.	
DR	PROSITE; PS00237; G.PROTEIN RECF_Fl_1; 1.	
DR	PROSITE; PS02623; G.PROTEIN RECF_Fl_2; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;	
KW	Lipoprotein; Palmitate.	
FT	SIGNAL 1 26	BY SIMILARITY.
FT	CHAIN 27 443	ENDOTHELIN B RECEPTOR.
FT	DOMAIN 27 102	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 103 127	1 (POTENTIAL).
FT	DOMAIN 128 133	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 139 164	2 (POTENTIAL).
FT	DOMAIN 165 176	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 177 198	3 (POTENTIAL).
FT	DOMAIN 199 219	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 220 244	4 (POTENTIAL).
FT	DOMAIN 245 272	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 273 297	5 (POTENTIAL).
FT	DOMAIN 298 325	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM 326 351	6 (POTENTIAL).
FT	DOMAIN 352 363	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 364 390	7 (POTENTIAL).
FT	DOMAIN 392 443	CYTOPLASMIC (POTENTIAL).
FT	DISULFID 175 256	BY SIMILARITY.
FT	CARBOHYD 60 60	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 354 354	N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	LIPID 403 403	PALMITATE (POTENTIAL).
FT	LIPID 404 404	PALMITATE (POTENTIAL).
FT	LIPID 406 406	PALMITATE (POTENTIAL).
SO	SEQUENCE 443 AA; 49594 MW; C3B066743990A3C7 CRC64;	

alignment scores:

```

222 CGACACTCTCCCGGGGAGAAGCAGTGCCCGAAGATTTTCTGAAGCCGGGG 271
    ||| ||| |||||::: ::: |||
5 ArgSerLeucyGlyArgAlaLeuValAlaLeuIlePhe..... 17
272 AAGCTGTGCAGCCGACGCCGCCGCCGCGAGCCCGGACACCGGCCA 321
    ::|||::: ::: ::: ||| ||| ||| |||
18 AlaCysGlyValAlaGlyValGlnSerGluGluArgGlyPhePro.Pro 33
322 CCTCC.....GCCCCACC 335
    ::: |||||
34 AlaGlyAlaThrProProAlaLeuArgThrGlyGluIleValAlaProPr 50
336 CACCCTCGCTTTCTCCGGCTCTCCCTGCCCGACGCGCGCGGACCCGG 385
    ||| ::||| |||||
50 oThrLysThrPhe.....TrpPro..... 56
386 CAGCTGTCTCGGCACGCCGAGCTCCAGGTCGAAAAAAGTGAAGGTGT 435
56 ..... 56
436 AAAAGCAGCACAGTGCATAAGAGATATTTCCCAAATTTGCCCTCAAGA 485
    ||| ::: |||||::: |||||::: |||||::: |||||:::
57 .....ArgGlySerAsnAlaSerLeuProArgSg 66
486 TGGAAACCCCT.....TTGCTCTAGGGGCATCCTTTTGCTGGCAGCTGGT 529

```

```
66 rSerSerProProGlnMetProLysGlyArgMetAla..... 79
530 GGATGTGTAATCAGTGATAATCCTCAGAGATACAGACAAATCTAAGCAA 579
79 ..... 79
580 TCATGTGGATGATTTCCACACATTTTCGGTGGCAGAGCTCAGCTTCCCTGG 629
79 ..... 79
630 TTACCACCTCATCAACCCACTAATTTGGTCTACCCAGCAATGGCTCAATG 679
80 .....GlyPro.ProAlaArgThrLeuThr 87
680 CACAACATATTGCCACAGCAGACTAAATTTACTTCAGCTTTCAAAATACAT 729
88 ProProProCysGluGlyProIleGluLeuLysAspThrPheLysTyrIle 104
730 TAACACTGTGATATCTTGATATTTTCATCGTGGGAATGGTGGGAATG 779
104 eAsnThrValSerCysLeuValPheValLeuGlyIleIleGlyAsnS 121
780 CAACCTCTGCTCAGGATCATTTACCAGAACAAATGTATGAGGAATGCCCC 829
121 erThrLeuLeuArgIleIleTyrLysAsnLysCysMetArgAsnGlyPro 137
830 AACGGCTGATAGCCAGCTTGTCCCTTGGACACCTTATCTATGTGTCAT 879
138 AsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleIleI 154
880 TGATCTCCCTATCAATGATTTAACTGCTGGCTGGCGCTGGCGCTTTTG 929
154 eAspIleProIleAsnValTyrLysLeuLeuAlaGluAspTrpPro.... 169
930 ATCAACATGACTTTCGGCTATTTCTTTCGACGCTGTTCCCTTTTGGCAG 979
170 .....PheGlyValGluMetCysLysLeuValProPheIleGln 182
980 AAGTCTCTGGTGGGATCACCCTCTCAACCTCTCGCTCTTAGTGTGA 1029
183 LysAlaSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAs 199
1030 CAGGTACAGACAGTGTGCTCTGAGTCGCTTCAGGAAATGGGATTC 1079
199 pArgTyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValP 216
1080 CTTTGTGTAACCTCCATTTGAAATGCTCCATCTGCTGCTGCTCTTATC 1129
216 rLysTrpThrAlaValGluIleValLeuIleTrpValValSerValVal 232
1130 CTGGCCATTCCTGAAGCATTGGCTTCCTCATGTTGACCTTTTGAATATAG 1179
233 LeuAlaValProGluAlaLeuGlyPheAspMetIleThrAspTyrLy 249
1180 GGTGAACAGCATAAACCCTGATCTCAATGCCACATCAAA.....T 1223
249 sGlyAsnArgLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaP 266
1224 TCATGAGCTTCTACCAAGATGTAAGGACTGGTGGCTCTTCGGGTCTAT 1273
266 hMetGlnPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyr 282
1274 TTCTGTATGCCCTTGGTGTGATCGCATCTTCTACACCTTCATCAGCTTG 1323
283 PheCysLeuProLeuAlaIleThrAlaPheTyrThrLeuMetThrCy 299
1324 TGAGATGTGAACAGAGAATGGCAGCTTGAGAATTCCTCAGTGAAC 1373
299 sGluMetLeuArgLysLysSerGly...MetGlnIleAlaLeuAsnAspH 315
1374 ATCTTAACGCGTGCAGAAAGTGGCAAAACAGTTTCTGCTGTTGTTGTA 1423
|||||
```

```
315 isLeuLysGlnArgArgGluValAlaLysThrValPheCysLeuValLeu 331
1424 ATTTTGTCTCTTTGCTGGTTCCTCTTCACTTAAGCCGTATATTGAAGAA 1473
332 ValPheAlaLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLe 348
1474 AACTGTGTATACGAATGGACAAGACCGATGTGAATTTACTTACTTCT 1523
348 uThrLeuTyrAspGlnAsnAspSerAsnArgCysGluLeuLeuSerPheL 365
1524 TACTCTCATGATACATCGGTATTAACCTTGGCAACCATGAATCATGT 1573
365 euLeuValLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCys 381
1574 ATAAACCCCATAGCTCTGTATTTTGTGACAAAGAAATTTAAAAATGTTT 1623
382 IleAsnProIleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPh 398
1624 CAGTCATGCTCTCTGCTGCTGTACACAGTCCAAAGCTCTGATGACCT 1673
398 eLysSerCysLeuCysCysTrpCys...GlnSerPheGluGlyLysGlnS 414
1674 CGTCCCATGAAACGGAACATCCAGTGGGAAGAACCCACGATCAAAAC 1723
414 erLeuGluLysGlnSerCysLeuLysPheLysAlaAsnAsp...His 429
1724 AACCAACAACACAGCCGAGCAGCAGCCATAGGACAGC 1759
430 GlyTyrAspAsnPheArgSerSerAsnLysTyrSer 441
seq_name: SwissProt_40:ETBR_COTJA
seq_documentation_block:
ID ETBR_COTJA STANDARD; PRT; 347 AA.
AC Q90328;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Endothelin B receptor (ET-B) (Endothelin receptor non-selective type)
DE (Fragment).
GN EDNRB.
OS Coturnix coturnix japonica (Japanese quail).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Coturnix.
OX NCBI_TaxID=93934;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96382520; PubMed=8790384;
RA Natarf V., Lecoin L., Eichmann A., le Douarin N.M.;
RT "Endothelin-B receptor is expressed by neural crest cells in the
RL avian embryo."
RL Proc. Natl. Acad. Sci. U.S.A. 93:9645-9650(1996).
CC -1- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE A
CC PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM (BY
CC SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X99295; CAA67681.1; -
CC GCRDB; GCR_1202; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
DR
```


RN (2)
 RP SEQUENCE FROM N.A.
 RX MEDLINE-91298956; PubMed-1648908;
 RA Ogawa Y., Nakao K., Arai H., Nakagawa O., Hosoda K., Suga S.,
 RA Nakanishi S., Imura H.;
 RT "Molecular cloning of a non-isopeptide-selective human endothelin
 RT receptor";
 RL Biochem. Biophys. Res. Commun. 178:248-255(1991).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93180293; PubMed-1291713;
 RA Arai H., Nakao K., Hosoda K., Ogawa Y., Nakagawa O., Komatsu Y.,
 RA Imura H.;
 RT "Molecular cloning of human endothelin receptors and their expression
 RT in vascular endothelial cells and smooth muscle cells";
 RL Jpn. Circ. J. 56:1303-1307(1992).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93155196; PubMed-8429023;
 RA Arai H., Nakao K., Takaya K., Hosoda K., Ogawa Y., Nakanishi S.,
 RA Imura H.;
 RT "The human endothelin-B receptor gene. Structural organization and
 RT chromosomal assignment";
 RL J. Biol. Chem. 268:3463-3470(1993).
 RN [5]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-9315496; PubMed-1713452;
 RA Sakamoto A., Yanagisawa M., Sakurai T., Takuwa Y., Yanagisawa H.,
 RA Masaki T.;
 RT "Cloning and functional expression of human cDNA for the ETB
 RT endothelin receptor";
 RL Biochem. Biophys. Res. Commun. 178:656-663(1991).
 RN [6]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93179382; PubMed-8440682;
 RA Elshourbagy N.A., Korman D.R., Wu H.L., Sylvester D.R., Lee J.A.,
 RA Nuthalaganti P., Bergsma D.J., Kumar C.S., Nambi P.;
 RT "Molecular characterization and regulation of the human endothelin
 RT receptors";
 RL J. Biol. Chem. 268:3873-3879(1993).
 RN [7]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LUNG;
 RX MEDLINE-93132955; PubMed-1282938;
 RA Haendler B., Hechler U., Schleuning W.D.;
 RT "Molecular cloning of human endothelin (ET) receptors ETA and ETB";
 RL J. Cardiovasc. Pharmacol. 20:S1-S4(1992).
 RN [8]
 RP SEQUENCE FROM N.A. (ISOFORM B).
 RC TISSUE-Placenta;
 RA Elshourbagy N., Adamou J., Gagnon A., Wu H., Pullen M., Nambi P.;
 RT "Molecular characterization of a human endothelin receptor splice
 RT variant";
 RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE FROM N.A. (ISOFORM DELTA 3).
 RX MEDLINE-99173868; PubMed-10072757;
 RA Tsutsumi M., Liang G., Jones P.A.;
 RT "Novel endothelin B receptor transcripts with the potential of
 RT generating a new receptor";
 RL Gene 228:43-49(1999).
 RN [10]
 RP PALMITOYLATION.
 RX MEDLINE-97407956; PubMed-9261180;
 RA Okamoto Y., Ninomiya H., Tanioka M., Sakamoto A., Miwa S., Masaki T.;
 RT "Palmitoylation of human endothelinB. Its critical role in G protein
 RT coupling and a differential requirement for the cytoplasmic tail by G
 RT protein subtypes";
 RL J. Biol. Chem. 272:21589-21596(1997).
 RN [11]
 RP REVIEW ON VARIANTS
 RX MEDLINE-98023959; PubMed-9359036;
 RA Hofstra R.M.W., Osinga J., Buys C.H.C.M.;
 RT "Mutations in Hirschsprung disease: when does a mutation contribute to
 RT the phenotype";
 RL Eur. J. Hum. Genet. 5:180-185(1997).
 RN [12]
 RP VARIANT HSCR CYS-276.
 RX MEDLINE-95094300; PubMed-8001158;
 RA Puffenberger E.G., Hosoda K., Washington S.S., Nakao K., Dewit D.,
 RA Yanagisawa M., Chakravarti A.;
 RT "A missense mutation of the endothelin-B receptor gene in multigenic
 RT Hirschsprung's disease";
 RL Cell 79:1257-1266(1994).
 RN [13]
 RP VARIANT WS/HSCR GLY-183.
 RX MEDLINE-96177681; PubMed-8634719;
 RA Attie T., Till M., Pellet A., Amiel J., Edery P., Boutrand L.,
 RA Munnich A., Lyonnet S.;
 RT "Mutation of the endothelin-receptor B gene in Waardenburg-
 RT Hirschsprung disease";
 RL Hum. Mol. Genet. 4:2407-2409(1995).
 RN [14]
 RP VARIANT HSCR ASN-305.
 RX MEDLINE-97005360; PubMed-8852659;
 RA Auricchio A., Casari G., Stalano A., Ballabio A.;
 RT "Endothelin-B receptor mutations in patients with isolated
 RT Hirschsprung disease from a non-inbred population";
 RL Hum. Mol. Genet. 5:351-354(1996).
 RN [15]
 RP VARIANTS HSCR SER-57; TRP-319 AND LEU-383.
 RX MEDLINE-97005361; PubMed-8852660;
 RA Amiel J., Attie T., Jan D., Pellet A., Edery P., Bidaud C., Lacombe D.,
 RA Tam P., Simeoni J., Flori E., Nihoul-Fekete C., Munnich A.,
 RA Lyonnet S.;
 RT "Heterozygous endothelin receptor B (EDNRB) mutations in isolated
 RT Hirschsprung disease";
 RL Hum. Mol. Genet. 5:355-357(1996).
 RN [16]
 RP VARIANT HSCR ILE-374, AND VARIANT SER-57.
 RX MEDLINE-96224410; PubMed-8630503;
 RA Hofstra R.M.W., Tan-Sindhunata G., Wu Y., Kamsteeg E.-J., Stulp R.P.,
 RA van Ravenswaaij-Arts C., Majoor-Krakauer D., Angrist M.,
 RA Chakravarti A., Meijers C., Buys C.H.C.M.;
 RT "A homozygous mutation in the endothelin-3 gene associated with a
 RT combined Waardenburg type 2 and Hirschsprung phenotype (Shah-
 RT Waardenburg syndrome)";
 RL Nat. Genet. 12:445-447(1996).
 RN [17]
 RP VARIANT HSCR SER-57.
 RX MEDLINE-98430996; PubMed-9760196;
 RA Svensson P.J., Anvret M., Molander M.L., Nordenskjold A.;
 RT "Phenotypic variation in a family with mutations in two Hirschsprung-
 RT related genes (RET and endothelin receptor B)";
 RL Hum. Genet. 103:145-148(1998).
 CC -I- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
 CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND DELTA;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- DISEASE: DEFECTS IN EDNRB ARE A CAUSE OF TYPE IV (WS4 OR SHAH-
 CC WAARDENBURG SYNDROME) (WS/HSCR) WHICH IS CHARACTERIZED BY THE
 CC ASSOCIATION OF WS AND HIRSCHSPRUNG DISEASE (HSCR).
 CC -I- DISEASE: DEFECTS IN EDNRB ARE THE CAUSE OF TYPE 2 HIRSCHSPRUNG
 CC DISEASE (HSCR2) (OR ANGLIOLIC MEGACOLON), A CONGENITAL DISORDER
 CC CHARACTERIZED BY ABSENCE OF ENTERIC GANGLIA ALONG A VARIABLE
 CC LENGTH OF THE INTESTINE. HSCR IS THE MOST COMMON CAUSE OF
 CC CONGENITAL INTESTINAL OBSTRUCTION EARLY SYMPTOMS RANGE FROM
 CC COMPLETE ACUTE NEONATAL OBSTRUCTION, CHARACTERIZED BY VOMITING,
 CC ABDOMINAL DISTENTION AND FAILURE TO PASS STOOL, TO CHRONIC
 CC CONSTIPATION IN THE OLDER CHILD.
 CC -I- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

FT	LIPID	405	405	PALMITATE (PO
1	1	1	1	1
2	2	2	2	2
3	3	3	3	3
4	4	4	4	4
5	5	5	5	5
6	6	6	6	6
7	7	7	7	7
8	8	8	8	8
9	9	9	9	9
10	10	10	10	10
11	11	11	11	11
12	12	12	12	12
13	13	13	13	13
14	14	14	14	14
15	15	15	15	15
16	16	16	16	16
17	17	17	17	17
18	18	18	18	18
19	19	19	19	19
20	20	20	20	20
21	21	21	21	21
22	22	22	22	22
23	23	23	23	23
24	24	24	24	24
25	25	25	25	25
26	26	26	26	26
27	27	27	27	27
28	28	28	28	28
29	29	29	29	29
30	30	30	30	30
31	31	31	31	31
32	32	32	32	32
33	33	33	33	33
34	34	34	34	34
35	35	35	35	35
36	36	36	36	36
37	37	37	37	37
38	38	38	38	38
39	39	39	39	39
40	40	40	40	40
41	41	41	41	41
42	42	42	42	42
43	43	43	43	43
44	44	44	44	44
45	45	45	45	45
46	46	46	46	46
47	47	47	47	47
48	48	48	48	48
49	49	49	49	49
50	50	50	50	50
51	51	51	51	51
52	52	52	52	52
53	53	53	53	53
54	54	54	54	54
55	55	55	55	55
56	56	56	56	56
57	57	57	57	57
58	58	58	58	58
59	59	59	59	59
60	60	60	60	60
61	61	61	61	61
62	62	62	62	62
63	63	63	63	63
64	64	64	64	64
65	65	65	65	65
66	66	66	66	66
67	67	67	67	67
68	68	68	68	68
69	69	69	69	69
70	70	70	70	70
71	71	71	71	71
72	72	72	72	72
73	73	73	73	73
74	74	74	74	74
75	75	75	75	75
76	76	76	76	76
77	77	77	77	77
78	78	78	78	78
79	79	79	79	79
80	80	80	80	80
81	81	81	81	81
82	82	82	82	82
83	83	83	83	83
84	84	84	84	84
85	85	85	85	85
86	86	86	86	86
87	87	87	87</	

CC entities requires a I

|||||

CC entities requires a I

```
156 ProfileAsnThrTyrLysLeuLeuAlaGluAspTrpPro..... 168
937 TGACTTGGCGGTATTTCTTTCGCAAGCTGTTCCTTTTTCGACAGTCCT 986
169 .....PheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAlaS 184
987 CGGTGGGATCAGCGTCTTCAACCTCTCCCTCTTCTAGTGTTCAGAGTAC 1036
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
1037 AGAGCAGTTCCTCCCTCGGTGCTGTCAGGGAATTCGGATTCCTTGGT 1086
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
1087 AACTGCCATTAATTCCTCCATCTGATCGTCCTCTTATCTCGCCCA 1136
217 pThrAlaValAlaGluIleValLeuIleTrpValValSerValLeuAlav 234
1137 TTCTCTGAAGGATTCGCTTCGTATGCTGATCCCTTTCGATATAGGGTGA 1186
234 alProGluAlaIleGlyPheAspMetIleThrSerAspTyrLysGlyLys 250
1187 CAGCATAAAACCTGTATGCTCAATGCCACATCAAAA.....TTCATGGA 1230
251 ProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMetG 267
1231 GTTCTACCAAGATGTAAGGACTGGTGGCTTCCTCGGTCTTATTTCTGTA 1280
267 nPheTyrLysThrAlaLysAspTrpLysPheSerPheTyrPheCysL 284
1281 TGCCCTGTGTCATCGCATCTTCTACACCTCATGACTTGTGAGATG 1330
284 euProLeuAlaIleThrAlaValPheTyrThrLeuMetThrCysGluMet 300
1331 TTGAACAGAGGAATGGCAGCTTCAGAAATTCGCTTCAGTGAACATCTTAA 1380
301 LeuArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuL 316
1381 CGAGCTCGAGAGTGGCAAAACAGTTTCTGCTGGTGTGTAATTTTG 1430
316 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 333
1431 CTCTTTGCTGGTTCCTCTCACTTAAGCGGTATATTGAGAAACATGTC 1480
333 lalaucysirpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 349
1481 TATAAGCAATGGCAAGAACCATGCTGAATTACTTACTTCTTACTGCT 1530
350 TyrAspGlnSerAsnProHisArgCysGluLeuLeuSerPheLeuLeuVa 366
1531 CATGGATTACATGGGTATTAACTTGGCAACCATGAATTCATGTATAAAC 1580
366 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnP 383
1581 CCATAGCTCTGATTTTTCGACCAAGAATTTTAAATTTTCTTCCAGTCA 1630
383 rolleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSer 399
1631 TGCTCTGCTGCTGCTTACAGTCCAAAGTCTCATGACCTCGGTCCC 1680
400 CysLeuCysCysTrpCys...GlnThrPheGluGluLysGlnSerLeuG 415
1681 CATGAACGGAACAAGCATCCAGTGGGAAGAACCCAGCATCAAAACACACA 1730
415 uGluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrA 431
1731 ACACAGACCGGAGCAGCCATAGGACAGC 1759
431 spAsnPheArgSerSerAsnLysTyrSer 440
```

seq_name: SwissProt_40:ETBR_CANFA

seq_documentation_block:

```
ID ETBR_CANFA STANDARD; PRT; 426 AA.
AC P56497;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
GN Non-selective type) (Fragment).
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RA Zemke D., Yuzbasivan-Gurkan V.;
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
CC FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC SUBCELLULAR LOCATION: Integral membrane protein.
CC SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; AF034530; AAC26970.1;
CC GCRdb; GCR_2495;
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODPSN.
CC PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
CC PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Lipoprotein; Palmitate.
FT SIGNAL 1 26
FT CHAIN 27 >426
FT DOMAIN 27 101
FT TRANSMEM 102 126
FT DOMAIN 127 137
FT TRANSMEM 138 163
FT DOMAIN 164 175
FT TRANSMEM 176 197
FT DOMAIN 198 218
FT TRANSMEM 219 243
FT DOMAIN 244 271
FT TRANSMEM 272 296
FT DOMAIN 297 324
FT TRANSMEM 325 350
FT DOMAIN 351 362
FT TRANSMEM 363 389
FT DOMAIN 390 >426
FT DISULFD 174 255
FT CARBOHYD 59 59
FT LIPID 402 402
FT LIPID 403 403
FT LIPID 405 405
FT NON_TER 426 426
SQ SEQUENCE 426 AA; 47876 MW; C5AF1675C30F1F26 CRC64;
```

alignment_scores:

Quality: 1228.00

Length: 322

Ratio: 4.264

Gaps: 3

Percent Similarity: 89.441

Percent Identity: 67.391

alignment_block:

US-09-931-157-1 x ETBR_CANFA


```
FT TRANSMEM 103 127 1 (POTENTIAL).
FT DOMAIN 128 138 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 139 164 2 (POTENTIAL).
FT DOMAIN 165 176 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 177 198 3 (POTENTIAL).
FT DOMAIN 199 219 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 220 244 4 (POTENTIAL).
FT DOMAIN 245 272 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 273 297 5 (POTENTIAL).
FT DOMAIN 298 325 6 (POTENTIAL).
FT TRANSMEM 326 351 7 (POTENTIAL).
FT DOMAIN 352 390 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 391 443 7 (POTENTIAL).
FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT DISULFID 175 256 BY SIMILARITY.
FT LIPID 403 403 PALMITATE (POTENTIAL).
FT LIPID 404 404 PALMITATE (POTENTIAL).
FT LIPID 406 406 PALMITATE (POTENTIAL).
FT VARIANT 126 126 I -> K (IN OLWS).
FT CONFLICT 11 11 V -> A (IN REF. 2).
FT CONFLICT 43 44 FE -> SG (IN REF. 2).
FT CONFLICT 64 66 PRL -> QRS (IN REF. 2).
FT CONFLICT 79 79 A -> E (IN REF. 2).
FT CONFLICT 92 92 E -> D (IN REF. 2).
FT CONFLICT 297 297 E -> M (IN REF. 2).
FT CONFLICT 369 369 E -> D (IN REF. 2).
FT CONFLICT 400 400 W -> S (IN REF. 2).
FT SEQUENCE 443 AA: 50006 MW: 74381DDFCF9A9AF4 CRC64;
```

alignment_scores:

```
Quality: 1226.00 Length: 359
Ratio: 3.955 Gaps: 5
Percent Similarity: 86.351 Percent Identity: 62.396
```

alignment_block:

US-09-931-157-1 x ETBR_HORSE

Align seg 1/1 to: ETBR_HORSE from: 1 to: 443

```
689 TGCCACAGCAGACTAAATTTACTTTCAGCTTTCAATACATTAACACTGT 738
|||||
91 CysGluArgThrIleGluIleLysGluThrPheLysThrIleAsnThrVa 107
739 GATATCTGTACTATTTTCATCGTGGGAATGTTGGGAATGCAACTCTGC 788
|||||
107 lValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeuL 124
789 TCAGGATCATTTACCAAGAAATGTATGAGGAATGCCCAACGCGCTG 838
|||||
124 euArgIleIleTyrlLysAsnLysCysMetArgAsnGlyProAsnIleLeu 140
839 ATAGCCAGTCTGCCCTTGAGACCTTATCTATGCTGCTCATTTGATCTCC 888
|||||
141 lIleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleIleIleAspIlePr 157
889 TATCAATGTATTTAAGCTGCTGCTGGCGCTGGCGCTTTTGTATCACAATG 938
|||||
157 oIleAsnValTyrlLysLeuLeuAlaGluAspTrpPro..... 169
939 ACTTTGCGGTPATTTCTTCCAAGCTGTCCCTTTTTCGAGAAGTCTCG 988
|||||
170 ..PheGlyValGluMetCysLysLeuValProPheIleGlnLysAlaSer 185
989 GTGGGATCACCTCTCAACCTCTGGCTCTTAGTGTTCACAGGTACAG 1038
|||||
186 ValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyAr 202
1039 AGCAGTTGCTCTCTGAGTCTGTTTCAGGAATTTGGGATTCCTTTGGTAA 1088
|||||
202 gAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpT 219
```

```
1089 CTGCCATTGAATTTGCTCCATCTGGATCTGCTCTTTATCTCTGCGCAT 1138
|||||
219 hrAlaValGluIleValLeuIleTrpValValSerValLeuAlaVal 235
1139 CTGAAACGATTTGCTTCTGCTGATGCTGCTTTCGAATATAGGGTGAACA 1188
|||||
236 ProGluAlaValGlyPheAspMetIleThrAlaAspTyrlLysGlySerTy 252
1189 GCATAAACCTGTATGCTCAATGCCACATCAAAA.....TTCATGAGT 1232
252 rLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMetGlnP 269
1233 TCTACCAAGATGAAAGACTGGTGGCTCTTCGGGTTCTTATTTCTGTATG 1282
269 heTyrlLysAsnAlaLysAspTrpTrpLeuPheSerPheTyrlPheCysLeu 285
1283 CCCTTGTGTGCTGCTGCTGCTTCTTACACCTCATGACTTGTGAGATGTT 1332
286 ProLeuAlaIleThrAlaPhePheTyrlThrLeuGluThrCysGluMetLe 302
1333 GAACAGAAGCAATGGCAGCTTGAGAATGCCCTCAGTGAACATCTTAAGC 1382
302 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 318
1383 AGGCTCGAGAAGTGGCAAAACAGTTTCTGCTTGGTTGTAATTTTGTCT 1432
318 lNArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 334
1433 CTTTGTGCTGCTCTCTTCACTTAAGCCGTATATTGAAGAAACCTGTGTA 1482
335 LeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysHisThrLeuTy 351
1483 TAACGAAATGGCAAGAACCGATGCTGAATTAATTAATTTCTTCTTACTGCTCA 1532
351 rAspGlnAsnAspProIleArgCysGluLeuLeuLeuSerPheLeuValL 368
1533 TGGATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAAACCCC 1582
368 euGluTyrlIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 384
1583 ATAGCTCTGTATTTTGTGACCAAGAAATTTAAATAATTTTTCAGTCTG 1632
385 lIleAlaLeuTyrlLeuValSerLysArgPheLysAsnCysPheLysTrpCy 401
1633 CCTCTGCTGCTGCTGTTTACCAGTCCAAAGCTCTGATGACCTCGTCCCA 1682
401 sLeuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGluA 417
1683 TGAACGGAAACAGCATCCAGTGGGAAGAACCCAGCATCAAAACACCAAC 1732
417 spLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrlAsp 432
1733 ACAGACCGGAGCAGCCATTAAGGACAGC 1759
433 AsnPheArgSerSerAsnLysTyrlSer 441
```

seq_name: SwissProt_40:ETBR_BOVIN

seq_documentation_block:

ID ETBR_BOVIN STANDARD; PRT: 441 AA.

AC P28088;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Endothelin B receptor precursor (ET-B) (endothelin receptor

DE Non-selective type).

GN EDNRB.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea;

OC Bovidae; Bovinae; Bos.

OX NCBI_TaxID=9913;

RN [1]

RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE-92078223; PubMed-1660473;
 RA Saito Y., Mizuno T., Itakura M., Suzuki Y., Ito T., Hagiwara H.,
 RA Hirose S.;
 RT "Primary structure of bovine endothelin ETB receptor and
 RT identification of signal peptidase and metal proteinase cleavage
 RT sites."; J. Biol. Chem. 266:23433-23437(1991).
 RL J. Biol. Chem. 266:23433-23437(1991).
 RN [2]
 RP SEQUENCE OF 124-127; 262-269; 304-315; 417-421 AND 424-432.
 RC TISSUE=Lung;
 RX MEDLINE-91358493; PubMed-1653249;
 RA Koruka M., Ito T., Hirose S., Lodhi K.M., Hagiwara H.;
 RT "Purification and characterization of bovine lung endothelin
 RT receptor."; J. Biol. Chem. 266:16892-16896(1991).
 RL J. Biol. Chem. 266:16892-16896(1991).
 RN [3]
 RP POST-TRANSLATIONAL MODIFICATIONS.
 RC TISSUE=Lung;
 RX MEDLINE-98086265; PubMed-9422751;
 RA Roos M., Soskic V., Poznanovic S., Godovac-zimmermann J.;
 RT "Post-translational modifications of endothelin receptor B from
 RT bovine lungs analyzed by mass spectrometry."; J.
 RL J. Biol. Chem. 273:924-931(1998).
 CC -1- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
 CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; D10994; BAA01762.1; -
 CC EMBL; D10989; BAA01762.1; JOINED.
 CC EMBL; D10990; BAA01762.1; JOINED.
 CC EMBL; D10991; BAA01762.1; JOINED.
 CC EMBL; D10992; BAA01762.1; JOINED.
 CC EMBL; D10993; BAA01762.1; JOINED.
 CC EMBL; D90456; BAA14422.1; -
 CC PIR; A41591; A41591.
 CC GCRdb; GCR_0237; -
 CC InterPro; IPR000276; GPCR_Rhodopsn.
 CC Pfam; PF00001; 7tm_1; 1.
 CC PRINTS; PR00237; GPCR_Rhodopsn.
 CC PROSITE; PS00237; G_PROTEIN_RECP_FL1_1; 1.
 CC PROSITE; PS0262; G_PROTEIN_RECP_FL2_1; 1.
 CC G-protein coupled receptor; Transmembrane; Signal; Lipoprotein;
 CC Palmitate; Phosphorylation.
 CC SIGNAL 1 26
 CC CHAIN 27 441 ENDOTHELIN B RECEPTOR.
 CC DOMAIN 27 100 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 101 125 1 (POTENTIAL).
 CC DOMAIN 126 136 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 137 162 2 (POTENTIAL).
 CC DOMAIN 163 174 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 175 196 3 (POTENTIAL).
 CC DOMAIN 197 217 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 218 242 4 (POTENTIAL).
 CC DOMAIN 243 270 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 271 295 5 (POTENTIAL).
 CC DOMAIN 296 323 CYTOPLASMIC (POTENTIAL).
 CC TRANSMEM 324 349 6 (POTENTIAL).
 CC DOMAIN 350 361 EXTRACELLULAR (POTENTIAL).
 CC TRANSMEM 362 388 7 (POTENTIAL).
 CC DOMAIN 389 441 CYTOPLASMIC (POTENTIAL).
 CC DISULFID 173 254 BY SIMILARITY.
 CC LIPID 402 402 PALMITATE.

FT LIPID 404 404 PALMITATE.
 FT MOD_RES 304 304 PHOSPHORYLATION.
 FT MOD_RES 418 418 PHOSPHORYLATION.
 FT MOD_RES 434 434 PHOSPHORYLATION (OR 435).
 FT MOD_RES 438 438 PHOSPHORYLATION.
 FT MOD_RES 439 439 PHOSPHORYLATION.
 FT MOD_RES 440 440 PHOSPHORYLATION.
 FT MOD_RES 441 441 PHOSPHORYLATION.
 SQ SEQUENCE 441 AA; 49371 MW; F634462C544DB0D2 CRC64;
 alignment_scores:
 Quality: 1221.50 Length: 376
 Ratio: 3.866 Gaps: 6
 Percent Similarity: 84.043 Percent Identity: 60.106
 alignment_block:
 US-09-931-157-1 x ETBR_BOVIN ..
 Align seg 1/1 to: ETBR_BOVIN from: 1 to: 441
 659 CTACCCAGCAATGGCTCAATGCACAACTAT..... 688
 72 ILeProArgGlyGlyArgMetAlaGlyLeProProArgThrProProPr 88
 689 .TGCCACAGCAGACTAAATTAATCTCAGCTTTCAAAATACATAACACTG 737
 88 oCysAspGlyProIleGluLeuLeuGlyGluThrPhelLysTyrIleAsnThrV 105
 738 TGATATCTTCTACTATTTTTCATGCTGGGAATGGGGGAATGCAACTCTG 787
 105 alValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeu 121
 788 CTGAGGATCATTTACCAAGACAATGTATGAGGAATGCCCAACGGCCT 837
 122 LeuArgIleIleTyrLysAsnLysCysMetArgAsnGlyProAsnIleLe 138
 838 GATAGGCAGCTTGGCCCTTGAGACCTTATCTATGCTGCTATTGATCATCC 887
 138 uIleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleIleIleAspIleP 155
 888 CTATCAATGATTTAAGCTGCTGGCTGGCGCTGGCCTTTTGTATCACAAT 937
 155 rIleAsnThrTyrLysLeuLeuAlaLysAspTrpPro..... 167
 938 GACTTTGGCGTATTCTTCTTCAAGCTCTCCCTTTTTCGAGAGAGTCTC 987
 168 ...PheGlyValGluMetCysLysLeuValProPheIleGlnLysAlaSe 183
 988 GGTGGGATCACCGTCTCAACCTCTCGCTCTTAGTGTGACAGGTACA 1037
 183 rValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrA 200
 1038 GACGAGTTGGCTCTCGAGTCGTGCTTTCAGGAATGGGATTCCTTTGGTA 1087
 200 rGAlaValAlaSerTrpSerArgIleGlyGlyIleGlyValProLysTrp 216
 1088 ACTGCCATTGAAATGTCTCATCTGCATCTGCTCTTTATTCCTGGCCAT 1137
 217 ThrAlaValGluLeuValLeuIleTrpValValSerValValLeuAlaVa 233
 1138 TCCTGAAGCGATTGGCTTCGTCATCGGTACCTTTTGAATATAGGGGTGAAC 1187
 233 lProGluAlaValGlyPheAspIleIleThrSerAspHisIleGlyAsnL 250
 1188 AGCATAAAACCTGATGCTCAATGCCACATCAAAA.....TTCATGGAG 1231
 250 ysLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMetGln 266
 1232 TTCTACCAAGATGTAAGGACTGGTGGCTCTCGGGTCTCTATTTCTCTAT 1281
 267 PheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheCysLe 283

1282 GCCCTTGGTGGCTGACCTGCGATCTTCTACACCTCATGCTGACATGT 1331
283 uProLeuAlaIleThrAlaLeuPheThrLeuMetThrCysGluMetL 300
1332 TGAACAGAAGGAATGCGACCTTGAGAAATGCCCTCAGTGAACATCTTAAG 1381
1382 euargLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLys 315
1382 CAGCGTCGAGAGTGGCAAAACAGTTTCTGCTGCTGGTGTAAATTTTC 1431
316 GlnArgGluValAlaLysThrValPheCysLeuValLeuValPheAl 332
1432 TCTTTCTGCTGCTCCCTCTTACCTTACCTTACCTTACCTTACCTT 1481
332 aleuCYstTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuT 349
1482 ATAACGAATGGACAAAGACCGATGGAATTAATTAATTAATTAATTA 1531
349 yrAspGlnHisAspProArgGlyCysGluPheLeuSerPheLeuVal 365
1532 ATGGATTACATCGGTATTAACCTTGGCAACCATGAATTCATGATAAACCC 1581
366 LeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPr 382
1582 CATAGCTCTGATTTTGTGAGCAAGAAATTAATAATTTCTTCCATCAT 1631
382 oileAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerC 399
1632 GCCTCTGCTGCTGCTTACCTTACCTTACCTTACCTTACCTTACCT 1681
399 ysLeuCYstTrpCys...GlnSerPheGluGluLysGlnSerLeuGlu 414
1682 ATGAACGGAACAGATCCAGTGGAGAGACCGATCAAAACACCAACAA 1731
415 GluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrAs 430
1732 CACAGACCGGAGCAGCATAAGACACAGC 1759
430 pAsnPheArgSerSerAsnLysTyrSer 439

seq_name: SwissProt_40:ET3R_XENLA

seq_documentation_block:

ID ET3R_XENLA STANDARD; PRT; 444 AA.

AC P32940;

DT 01-OCT-1993 (Rel. 27, Created)

DT 01-OCT-1993 (Rel. 27, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Endothelin-3 receptor precursor (ET-C).

OS Xenopus laevis (African clawed frog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC Xenopodinae; Xenopus.

OX NCBI_TaxID=8355;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Dermal melanophores;

RX MEDLINE=93366838; PubMed=8360195;

RA Karne S., Javawickreme C.K., Lerner M.R.;

RT "Cloning and characterization of an endothelin-3 specific receptor

RT (ETC receptor) from Xenopus laevis dermal melanophores.";

RL J. Biol. Chem. 268:19126-19133(1993).

CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-3. MEDIATES ITS ACTION BY

CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-

CC CALCIUM SECOND MESSENGER SYSTEM.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL; L20299; AAA49704.1; -

DR PIR; A48538; A48538.

DR GCRDB; GCR_0764; -

DR InterPro; IPR000276; GPCR_Rhodopsn.

DR Pfam; PF00001; 7cm1; 2.

DR PRINTS; PR00237; GPCRHHODOPSN.

DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; FALSE_NEG.

DR PROSITE; PS0262; G-PROTEIN_RECEP_F1_2; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.

FT SIGNAL 1 18

FT CHAIN 19 444

FT DOMAIN 19 88

FT TRANSMEM 89 113

FT DOMAIN 114 124

FT TRANSMEM 125 152

FT DOMAIN 153 161

FT TRANSMEM 162 180

FT DOMAIN 181 201

FT TRANSMEM 202 226

FT DOMAIN 227 254

FT TRANSMEM 255 279

FT DOMAIN 280 307

FT TRANSMEM 308 334

FT DOMAIN 335 351

FT TRANSMEM 360 386

FT DOMAIN 387 444

FT CARBOHYD 60 60

FT SEQUENCE 444 AA; 50172 MW; 80AB7EC36720D783 CRC64;

alignment_scores:

Quality: 1014.50

Length: 392

Ratio: 3.404

Gaps: 11

Percent Similarity: 76.020 Percent Identity: 52.041

alignment_block:

US-09-931-157-1 x ET3R_XENLA

Align seg 1/1 to: ET3R_XENLA from: 1 to: 444

665 AGCAATGGCTCAATGCACAACTAT.....TGCCC 693
|||||.....|

62 SerThrGlyAsnValLeuAsnMetSerProProProSerProCysLe 78
|||||.....|

694 ACAGCAGACTAAATTTACTTTCAGCTTCAATACATTAACTGTGATAT 743
.....|

78 userArgAlaLysIleArgHisAlaPheLysTyrValThrIleLeuS 95
.....|

744 CTTGTACTATTTCATCGTGGGAATGGTGGGAATGCAACTCTGCTCAGG 793
|||||.....|

95 erCysValIlePheLeuValGlyIleValGlyAsnSerThrLeuLeuArg 111
.....|

794 ATCATTTTACCAGAACAAATGTATGAGGAATGCCCAACGCGCTGATAGC 843
|||||.....|

112 IleIleTyrLysAsnLysCysMetArgAsnGlyProAsnValLeuIleAl 128
.....|

844 CAGCTTGGCCCTGGGAGACCTTATCTATCTGCTGCTGCTGCTGCTGCT 892
|||||.....|

128 aserLeuAlaLeuGlyAspLeuPheTyrIleLeuIleAlaIleProIleI 145
.....|

893AATGATTTTAACTGCTGGCTGGCGCTGGCGCTTTTGTATCAAT 937
|||||.....|

145 leSerIleSerPheTrpLeuSerThrGly.....HisSer 156
|||||.....|

938 GACTTTGGCGTATTCTTTTGCAGAGCTGTCCCTTTTGCAGAAAGTCCTC 987
.....|

157 GluTyr.....IleTyrGlnLeuValHisLeuTyrAr 167
.....|

988 GTGGGGATCACCGTCCTCAACTCTGCGCTCTAGTGTGACAGGTACA 1037
.....|

```
167 gAlaArgValTyrSerLeuSerLeuCysAlaLeuSerIleAspArgTyrA 184
1038 GAGCAGTTCCTCGTCGAGTCGCTTCCAGGAATGGATTCCTTGGTA 1087
|||||
184 rGAlaValAlaSerTrpAsnArgIleArgSerIleGlyIleProValArg 200
|||||
1088 ACTGCCATTGAATGCTCCATCGATCGATCGTCCTTTATPCTCGGCAT 1137
|||||
201 LysAlaIleGluLeuThrLeuIleTrpAlaValAlaIleIleValAla 217
|||||
1138 TCCTGAAGCATTGGCTCGTCATGGTACCTTTGAATATAGGGGTGAAC 1187
|||||
217 lProGluAlaIleAlaPheAsnLeuValGluLeuAspPheArgGlyGlnT 234
|||||
1188 AGCATAAACCTGTATCCTC.....AATGCCACATCAAAATTCATGGC 1231
|||||
234 hrIleLeuValCysMetLeuProMetGluGlnThrSerAspPheMetArg 250
|||||
1232 TTCTACCAAGATGTAAGGACTGCTGCTTCCTCGGTTCTATTCTGTAT 1281
|||||
251 PheTyrGlnGluValLysValTrpLeuPheGlyPheTyrPheCysLe 267
|||||
1282 GCCCTTGGTGGTACTCGCATCTTCTACACCTCATGACTGTGTGAGATGT 1331
|||||
267 uProLeuAlaCysThrGlyValPheTyrThrLeuMetSerCysGluMetL 284
|||||
1332 TGAACAGAGATGGCAGCTGTGAGATTCGCCCTCAGTGAACATCTTAAG 1381
|||||
284 euSerIleLysAsnGly...MetArgIleAlaLeuAsnAspHisMetLys 299
|||||
1382 CAGCGTCGAGAAGTGGCAAAACAGTTTCTGCTTGTGTGTAATTTTGC 1431
|||||
300 GlnArgArgGluValAlaLysThrValPheCysLeuValIlePheAl 316
|||||
1432 TCTTTGCTGCTTCCTCTTCACTTAAGCCGTATATG..... 1468
|||||
316 aLeuCysTrpLeuProLeuHisValSerSerIlePheValArgLeuSerA 333
|||||
1469 .....AAGAAACTGTGTATTAAC 1486
|||||
333 laThrValLysArgAlaCysIleLeuLysAsnLysArgSer.CysIleMe 349
|||||
1487 GAAATGGACAAGAACCCGATGTAATCTACTTACTTCTTACTGCTCATGGA 1536
|||||
349 taIaGluIleGlnThrGlyValAsnTyr..GlnLeuLeuMetValMetAs 365
|||||
1537 TTATATCGGTATTAACTTGGAACCATGAATTCATGTATATAAACCCCATAG 1586
|||||
365 nTyrThrGlyIleAsnMetAlaSerLeuAsnSerCysIleGlyProValA 382
|||||
1587 CTCTGTATTGTGTCAGCAAGAAATTTAAATTTGTTTCCAGTCATGCCTC 1636
|||||
382 laLeuTyrPheValSerArgLysPheLysAsnCysPheGlnSerCysLeu 398
|||||
1637 TGTGCTGCTGTTTACCAGTCCAAAAGTGTGATGACCTCGGTCGCCATGAA 1686
|||||
399 CysCysTrpCysHisArgProThrLeuThrIleThrProMetAspGluLys 415
|||||
1687 CGGAACAGCATTCGATGGAGAACACCGATCAAAACAAACAC.....A 1730
|||||
415 sGlySerGlyGlyLysTrpLys.....AlaAsnGlyHisAspLeuA 429
|||||
1731 ACACAGACCGGAGCAGCCATTAAG 1753
|||||
429 spLeuAspArgSerSerArg 436
|||||
seq_name: SwissProt_40-GRPR_MOUSE
seq_documentation_block:
ID GRPR_MOUSE STANDARD; PRT; 384 AA.
AC P21729;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
```

```
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Gastrin-releasing peptide receptor (GRP-R) (GRP-preferring bombesin
DE receptor).
GN GRPR.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC STRAIN=SWISS; TISSUE=Fibroblast;
RX MEDLINE=91110536; PubMed=1671171;
RA Battey J.F., Way J.M., Corjay M.H., Shapira H., Kusano K., Harkins R.,
RA Wu J.M., Slattery T., Mann E., Feldman R.I.;
RT "Molecular cloning of the bombesin/gastrin-releasing peptide receptor
RT from Swiss 3T3 cells.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:395-399(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=SWISS; TISSUE=Fibroblast;
RX MEDLINE=91187004; PubMed=1707129;
RA Spindel E.R., Giladi E., Brehm P., Goodman R.H., Segerson T.P.;
RT "Cloning and functional characterization of a complementary DNA
RT encoding the murine fibroblast bombesin/gastrin-releasing peptide
RT receptor.";
RL Mol. Endocrinol. 4:1956-1963(1990).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93305484; PubMed=8391296;
RA Giladi E., Nagalla S.R., Spindel E.R.;
RT "Molecular cloning and characterization of receptors for the
RT mammalian bombesin-like peptides.";
RL J. Mol. Neurosci. 4:41-54(1993).
CC -1- FUNCTION: RECEPTOR FOR GASTRIN RELEASING PEPTIDE (GRP). THIS
CC RECEPTOR MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT
CC ACTIVATE A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: BRAIN (HYPOTHALAMUS), PANCREATIC ACINAR CELLS,
CC AND FIBROBLASTS.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; M57922; AAA75650.1; -.
DR EMBL; M61000; AAA37744.1; -.
DR PIR; A36553; A36553.
DR PIR; A39003; A39003.
DR GCRDB; GCR_0096; -.
DR GCRDB; GCR_0097; -.
DR MGD; MGI:95836; Grpr.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRRHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein;
KW Lipoprotein; Palmitate.
FT DOMAIN 1 39 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 40 63 1 (POTENTIAL).
FT DOMAIN 64 77 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 78 97 2 (POTENTIAL).
FT DOMAIN 98 115 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 116 137 3 (POTENTIAL).
FT DOMAIN 138 153 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 154 175 4 (POTENTIAL).
FT DOMAIN 176 209 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 210 235 5 (POTENTIAL).
```


OM of: US-09-931-157-1 to: PIR_71.* out_format : pfs

Date: Oct 1, 2002 1:10 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

```
-MODEL=frame+np2.model -DEV=xlp  
-O/Cgnt2_1/USPTO_spool/US09931157/runat_30092002_l10421_l1282/app_query.fasta_1.8657  
-DB=PIR_71 -OFMT=fastan -SUFFIX=trp -GAPOP=12.000 -GAPEXT=4.000  
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -CGAPOP=4.500  
-CGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000  
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000  
-DELEX=7.000 -START=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0  
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500  
-MINLEN=0 -MAXLEN=200000000 -USER=US09931157 -CGNLS_1335  
-NCLPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30  
-NO_XLPXY -WAIT -THREADS=1
```

Search information block:

Query: US-09-931-157-1

Query length: 4105

Database: PIR_71.*

Database sequences: 283138

Database length: 96089334

Search time (sec): 273.170000

score_list:	Sequence	Strd Orig	25score	EScore	Len	Documentation
PIR2:A44158	+	2282.00	3626.07	1.6e-194	427	! endothelin receptor type A precursor
PIR2:S13424	+	2155.00	3423.01	3.3e-183	427	! endothelin receptor - bovine
PIR2:A04040	+	2103.50	3340.69	1.3e-178	426	! endothelin 1 and 2 receptor precursor
PIR2:AS4126	+	1623.00	2572.70	7.9e-136	415	! endothelin receptor-AX - Africa
PIR2:S13425	+	1240.00	1959.76	1.0e-101	441	! endothelin receptor ETB - rat
PIR1:JQ1042	+	1240.00	1959.74	1.0e-101	442	! endothelin receptor B precursor
PIR2:S17950	+	1240.00	1959.74	1.0e-101	442	! nonselective-type endothelin receptor
PIR2:A41591	+	1221.50	1930.18	4.6e-100	441	! endothelin-3 receptor precursor
PIR2:A48538	+	1014.50	1599.15	1.2e-81	444	! endothelin-1 receptor homolog -
PIR2:BS7508	+	605.00	927.27	2.7e-45	117	! bombesin/gastrin-releasing peptide
PIR2:A39003	+	465.00	721.98	1.0e-32	384	! bombesin/ GRP receptor - rat
PIR2:S17682	+	463.00	718.79	1.6e-32	384	! gastrin-releasing peptide receptor
PIR2:A41007	+	455.50	706.80	7.3e-32	390	! bombesin receptor, neuromedin-B
PIR2:JH0374	+	454.50	705.05	9.0e-32	390	! bombesin receptor, neuromedin-B
PIR2:S29480	+	422.00	652.87	7.1e-29	399	! bombesin receptor - guinea pig
PIR2:IS0102	+	411.50	636.65	6.0e-28	376	! Phe13 bombesin receptor - Bomb
PIR2:A46632	+	409.00	632.08	1.0e-27	399	! bombesin-like peptide receptor
PIR2:JC5501	+	394.50	604.76	2.2e-26	613	! endothelin receptor type-B-like
PIR2:S56517	+	323.50	495.40	4.2e-20	398	! mu opioid receptor - rat
PIR2:A57510	+	322.50	493.80	5.2e-20	398	! mu opioid receptor - mouse
PIR2:S56533	+	315.50	482.56	2.2e-19	400	! mu opiate receptor - human
PIR2:IA9022	+	315.00	482.59	2.4e-19	367	! kappa opioid receptor 3 - mouse
PIR2:JC2421	+	315.00	482.59	2.4e-19	367	! opioid receptor homolog, MOR-C
PIR2:S43087	+	313.50	480.12	3.2e-19	370	! orphan opioid receptor ORL1 - h
PIR2:S56504	+	313.50	479.41	3.3e-19	398	! mu opioid receptor - rat
PIR2:A48227	+	313.00	479.06	3.6e-19	380	! kappa opioid receptor 1 - mouse
PIR2:S56520	+	312.00	477.80	4.4e-19	367	! G protein-coupled receptor - rat
PIR2:S55693	+	312.00	477.16	4.4e-19	392	! opioid receptor mu variant MOR1
PIR2:S36143	+	310.00	474.26	6.6e-19	380	! kappa opioid receptor - rat
PIR2:JC2434	+	308.00	471.07	1.0e-18	380	! delta opioid receptor - mouse
PIR2:JE0087	+	305.00	466.45	1.8e-18	373	! delta opioid receptor - Zebraf
PIR2:JC2338	+	305.00	466.27	1.8e-18	380	! kappa opioid receptor 1 - human
PIR2:AS5259	+	302.00	461.47	3.4e-18	380	! kappa opioid receptor - guinea
PIR2:JC7677	+	299.00	455.64	6.5e-18	423	! allatostatin receptor - America
PIR2:T27866	+	298.00	453.86	8.0e-18	444	! hypothetical protein ZK455.3 -
PIR2:TC7209	+	297.50	453.93	8.7e-18	394	! galanin receptor - fruit fly (D
PIR2:A44081	+	294.00	447.27	1.8e-17	440	! kappa-type opioid receptor - hu
PIR2:DA1789	+	291.50	441.97	3.1e-17	504	! tachykinin receptor NKD - fruit
PIR2:A41795	+	290.50	443.37	3.6e-17	369	! somatostatin receptor 2 - mouse

```
PIR2:JM0605 + 290.50 442.89 3.6e-17 388 ! somatostatin receptor 4  
PIR2:A41738 + 289.00 439.08 5.1e-17 449 ! neuropeptide Y receptor  
PIR2:A41795 + 286.00 435.62 9.1e-17 391 ! somatostatin receptor 1  
PIR2:C41795 + 286.00 435.62 9.1e-17 391 ! somatostatin receptor 1  
PIR2:A47249 + 284.50 433.39 1.2e-16 384 ! brain-specific somatostat  
seq_name: pir2:A44158  
seq_documentation_block:  
seq_documentation_block:  
seq_documentation_block:  
N:Alternate names: endothelin 1 receptor; endothelin-A receptor  
C:Species: Homo sapiens (man)  
C>Date: 30-Sep-1993 #sequence.revision 30-Sep-1993 #text.change 21-Jul-2000  
C:Accession: A44158; S16980; J07972; JS0607; PC1245; A45481; S62869  
R:Hosoda, K.; Nakao, K.; Tamura, N.; Arai, H.; Ogawa, Y.; Suga, S.; Nakanishi, S.;  
J. Biol. Chem. 267, 18797-18804, 1992  
A:Title: Organization, structure, chromosomal assignment, and expression of the ge  
A:Reference number: A44158; MUID:92406798  
A:Accession: A44158  
A:Molecule type: DNA  
A:Residues: 1-427 <HOS>  
A:Cross-references: GB:D11145  
A:Note: sequence extracted from NCBI backbone (NCBIP:113475)  
R:Hosoda, K.; Nakao, K.; Arai, H.; Suga, S.; Ogawa, Y.; Mukoyama, M.; Shirakami, C  
FEBS Lett. 287, 23-26, 1991  
A:Title: Cloning and expression of human endothelin-1 receptor cDNA.  
A:Reference number: S16980; MUID:91348221  
A:Accession: S16980  
A:Molecule type: mRNA  
A:Residues: 1-427 <HO2>  
A:Cross-references: GB:X61950; NID:g288312; PIDN:CAA43953.1; PID:g288313  
R:Cyr, C.; Huebner, K.; Druck, T.; Kriis, R.  
Biochem. Biophys. Res. Commun. 181, 184-190, 1991  
A:Title: Cloning and chromosomal localization of a human endothelin eta receptor.  
A:Reference number: J07972; MUID:92068188  
A:Accession: J07972  
A:Molecule type: mRNA  
A:Residues: 1-427 <CYR>  
A:Cross-references: GB:S67127; NID:g239560; PIDN:AAB20407.1; PID:g239561  
A:Experimental source: placenta  
R:Adachi, M.; Yang, Y.Y.; Furuichi, Y.; Miyamoto, C.  
Biochem. Biophys. Res. Commun. 180, 1265-1272, 1991  
A:Title: Cloning and characterization of cDNA encoding human A-type endothelin rec  
A:Reference number: JS0607; MUID:92062153  
A:Accession: JS0607  
A:Molecule type: mRNA  
A:Residues: 1-427 <ADA>  
A:Cross-references: GB:S63938; NID:g238635; PIDN:AAB20278.1; PID:g238636  
R:Yang, H.; Tabuchi, H.; Furuichi, Y.; Miyamoto, C.  
Biochem. Biophys. Res. Commun. 190, 332-339, 1993  
A:Title: Molecular characterization of the 5'-flanking region of human genomic ET  
A:Reference number: PC1245; MUID:93151811  
A:Accession: PC1245  
A:Molecule type: DNA  
A:Residues: 1-140 <YAN>  
A:Cross-references: GB:S55772; NID:g264726; PIDN:AAB25212.1; PID:g264727  
R:Elshourbagy, N.A.; Korman, D.R.; Wu, H.L.; Sylvester, D.R.; Lee, J.A.; Nuthalaga  
J. Biol. Chem. 268, 3873-3879, 1993  
A:Title: Molecular characterization and regulation of the human endothelin recept  
A:Reference number: A45481; MUID:93179382  
A:Accession: A45481  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-427 <ELS>  
A:Experimental source: lung  
A:Note: sequence extracted from NCBI backbone (NCBIN:125752, NCBIP:125753)  
R:Miyamoto, Y.; Yoshimasa, T.; Arai, H.; Takaya, K.; Ogawa, Y.; Itoh, H.; Nakao, K  
Biochem. J. 313, 795-801, 1996  
A:Title: Alternative RNA splicing of the human endothelin-A receptor generates mul  
A:Reference number: S62869; MUID:96190719  
A:Accession: S62869  
A>Status: preliminary  
A:Molecule type: mRNA
```

A;Residues: 1-427 <MI>

A;Cross-references: GB:S81539; NID:gl1478475; PIDN:AAB36325.1; PID:gl1478476
C;Comment: This endothelin receptor, designated subtype A, binds preferentially to endothelin-1.
C;Genetics:
A;Gene: GDB:EDNRA
A;Cross-references: GDB:129028; OMIM:131243
A;Map position: 4pter-4qter
A;Introns: 140/3; 183/2; 249/3; 300/3; 345/2; 381/3
C;Superfamily: endothelin receptor B
C;Keywords: alternative splicing; G protein-coupled receptor; glycoprotein; lipoprotein; F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-427/Product: endothelin receptor A #status predicted <WAT>
F;81-102/Domain: transmembrane #status predicted <TM1>
F;113-132/Domain: transmembrane #status predicted <TM2>
F;160-181/Domain: transmembrane #status predicted <TM3>
F;206-229/Domain: transmembrane #status predicted <TM4>
F;257-278/Domain: transmembrane #status predicted <TM5>
F;307-328/Domain: transmembrane #status predicted <TM6>
F;348-372/Domain: transmembrane #status predicted <TM7>
F;389-393/Domain: transmembrane #status predicted <TM8>
F;289-373, 391, 421/Binding site: phosphate (Ser) (covalent) #status predicted
F;385-386, 387, 388/Binding site: palmitate (Cys) (covalent) #status predicted
F;417/Binding site: phosphate (Thr) (covalent) #status predicted

alignment_scores:

Quality: 2282.00 Length: 427
Ratio: 5.344 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-931-157-1 x A44158

Align seg 1/1 to: A44158 from: 1 to: 427

485 ATGGAACCTTTGGCTCAGGCGATCCTTTGGCTGGCGACTGGTGGATG 534
|||||
1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyc 17
535 TGTATCAGTATATCTCTGAGATACAGACAAATCTAAGCAATCATG 584
|||||
17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHlsv 34
585 TGGATGATTTCACCACTTTTCGTGGCAGAGCTCAGCTTCCTGGTTACC 634
|||||
34 alAspAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
635 ACTCATCAACCCACTAAATTTGGTCTCCACCAATGGCTCAATGCACAA 684
|||||
51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMethisAs 67
685 CTATTGCCACAGACAGACTAAATTTACTTCACGCTTCAATATACATA 734
|||||
67 nfyCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
735 CTGTGATATCTGTACTATTTTCATCGTGGGAATGGTGGGAATGCAACT 784
|||||
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
785 CTGCTCAGGATCATTTACCAACAATAATGTATGAGAAATGGCCCCAACGC 834
|||||
101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
835 GCTGATAGCAGCTTTGGCTGGAGACCTTATCTATGTTGGTCATGTATC 884
|||||
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValIleAspL 134
885 TCCCTATCAATCATTTAAGCTGCTGGCTGGCGCTGGCTTTGTATCAC 934
|||||
134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpPheAspHis 150
935 AATGACTTTGGCGTATTTCTTGCAGAGCTGTTCCCTTTTGCAGAGTC 984
|||||

151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysse 167
985 CTCGGTGGGATCACCGCTCAACCTCTGGCTCTTAGTGTGACAGGT 1034
|||||
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
1035 ACAGAGCAGTTGGCTCCCTGGAGTCTGGTTCAGGGAATGGGATTCCTTTG 1084
|||||
184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
1085 TAACTCGCAATGAAATTTCTCCATCTCGATCCCTGCTCTTTATCTCTGC 1134
|||||
201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
1135 CATTCCTGAAGCGATTTGGCTTCGTCATGCTACCTTTGAATATAGGGGTG 1184
|||||
217 arIleProGluAlaIleGlyPheValMetValProPheGluTyrArgGlyG 234
1185 AACAGCATAAACCTGATGCTCAATGCCACATCAAAATTCATGAGATTC 1234
|||||
234 IuGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
1235 TACCAAGATGTAAGAGACTGGTGGCTCTTCGGGTCTTATTTCTGTATGCC 1284
|||||
251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPr 267
1285 CTTGTGTGCTGCTGCTGCTCTTCTACACCTCATGACTTGTGAGATGTTGA 1334
|||||
267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
1335 ACAGAAGGAATGGCAGCTTGAGAAATGGCTCAGTCAACATCTTAAGCAG 1384
|||||
284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
1385 CQTCGAGAGTGGCAAAACAGTTTTCTGCTGGTTGTAATTTTGTCTCT 1434
|||||
301 ArgArgGluValAlaLysThrValPheCysLeuValIlePheAlaLe 317
1435 TTGCTGGTTCCCTCTCACTTAAGCGGTATATTGAAGAAACTGTGTATA 1484
|||||
317 ucysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 AGCAATGGACAGAACCGATGTGAATTAATTAATTTCTTACTGCTCATG 1534
|||||
334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAAACCCAT 1584
|||||
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProI 367
1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTTCCAGTCATGCC 1634
|||||
367 eaLeuLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TGTGCTGCTGCTTACAGTCCAAAGTCTGATGACCTCGGTCCCATG 1684
|||||
384 euCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValProMet 400
1685 AACGGAACAGCATCCAGTGGAGAGACCATCAAAACCAACCAACCAACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnHisAsnTh 417
1735 AGACCGGAGCAGCATTAAGGACAGCATGAAC 1765
|||||
417 raspArgSerSerHisLysAspSerMetAsn 427
seq_name: pir2.S13424

seq_documentation_block:

endothelin receptor - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 16-Jul-1999
C;Accession: S13424

R;Arai, H.; Hori, S.; Aramori, I.; Ohkubo, H.; Nakanishi, S.

Nature 348, 730-732, 1990

A;Title: Cloning and expression of a cDNA encoding an endothelin receptor.

A;Reference number: S13424; MUID:91080923

A;Accession: S13424

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-427 <ARA>

A;Cross-references: GB:X57765; NID:g121; PIDN:CAA0917.1; PID:g122

C;Superfamily: endothelin receptor B

alignment_scores:

Quality: 2155.00 Length: 427
Ratio: 5.143 Gaps: 0
Percent Similarity: 98.126 Percent Identity: 94.614

alignment_block:

US-09-931-157-1 x S13424

Align seg 1/1 to: S13424 from: 1 to: 427

```

485 ATGGAACCCCTTGGCTCAGGCATCCTTTGGCTGGCACTGGTTGGATG 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MetGluThrPheTrpLeuArgLeuSerPheTrpValAlaLeuValGlyCl 17

535 TGTATCATGATGATATCTCTGAGATACAGCACAAATCTAAGCAATCATG 584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
17 yValIleSerAspAsnProGluSerTy-SerThrAsnLeuSerIleHisv 34

585 TGGATGATTTTACCAGCTTTTCTGGCAGACAGCTCAGCTTCTGTTACC 634
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
34 aAspSerValAlaThrPheHisGlyThrGluLeuSerPheValValThr 50

635 ACTCATCAACCCACATAATTTGGTCTACCCAGCAATGGCTCAATGCACAA 684
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
51 ThrHisGlnProThrAsnLeuAlaLeuProSerAsnGlySerMetHisAS 67

685 CTATTGCCACAGCAGACATAAATTTACTTACAGCTTTTCAATATACATA 734
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
67 nTyrcysProGlnGlnThrLysIleThrSerAlaPheLysTyIleAsnt 84

735 CTGTATATCTGTACTATTTTCATCGTGGGAATGGTGGGAATGCAACT 784
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100

785 CTGCTCAGGATCATTTACCAACAATAATGTATGAGGAATGGCCCAACGC 834
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
101 LeuLeuArgIleIleTyGlnAsnLysCysMetArgAsnGlyProAsnAl 117

835 GCTGATAGCCAGCTCTTGGCCCTTGGAGACCTTATCTATGTGTCATTGATC 884
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyValValIleAspL 134

885 FCCCTATCAATGATTTAAGCTGCTGGCTGGCGCTGGCCCTTTTGCATCAC 934
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheGluGln 150

935 AATGACTTTGGCGTATTTCTTTTGAAGCTGTTCCCTTTTTCACGAAGTC 984
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
151 AsnAspPheGlyValPheLeucysLysLeuPheProPheLeuGlnLysSe 167

985 CTCGGTGGGGATCACCGCTCAACCTCTGGCTCTTACTGTGTGACAGGT 1034
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
167 rSerValGlyIleThrValLeuAsnLeuLysAlaLeuSerValAspArgT 184

1035 ACAGAGAGTGGCTTCCTGGAGTCTGTTACAGGAATGGGAATTCCTTTG 1084
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
184 yArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200

1085 GPAACCTGCATTAATTTGCTCCATCTGGATCCTCTTATCTCTGCTGCC 1134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217

```

```

1135 CATTCCTCAAGCGATTGGCTTGGTTCATGTTACCTTTGAATATAGGGTG 1184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
217 alleProGluAlaIleGlyPheValMetValProPheGluTyLysGlyA 234

1185 AACACATAAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
234 IaGlnHisArgThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250

1235 TACCAAGATGTAAGGACTGGTGGCTTCTCGGGTCTTATTTCTGTATGCC 1284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyPheCysMetPr 267

1285 CTTGTGTGCACTGCGATCTTCTACACCTCATGACTTGTGTGAGATGTTGA 1334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
267 oLeuValCysThrAlaIlePheTyThrLeuMetThrCysGluMetLeuA 284

1335 ACAGAAGGAATGGCAGCTTGAGAAATGCCCTCAGTGAACATCTTAAGCAG 1384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300

1385 GCTCAGAGAGTGGCAAAACAGTTTTCTGCTTGGTTGTAAATTTTGTCTCT 1434
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317

1435 TTGCTGTTCCCTCTTCACTTAAGCCGTATATTGAAGAAACTGTGTATA 1484
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyra 334

1485 ACGAATGGACAGAAGCAACGATGTAATTTACTTACTTCTTACTGCTCATG 1534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
334 spGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350

1535 GATTACATCGGTATTAACTTGGCAACCATGATTCATGTATTAACCCCAT 1584
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
351 AspTyrlleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367

1585 AGCTCTGTATTTTGTGACCAAGAAATTTAAATAATTTTCCAGTCATGCC 1634
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
367 eAlaLeuTyrrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384

1635 TCTGCTGCTGCTGTTTACCAGTCCAAAGTCTGATGACCTCGTCCCCATG 1684
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
384 euCysCysCysTyrrGlnSerLysSerLeuMetThrSerValProMet 400

1685 AACGAACAAGCATCCAGTGGGAAGAACACCATCAAAACACCAACACAC 1734
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
401 AsnGlyThrSerIleGlnTrpLysAsnHisGluGlnAsnAsnHisAsnTh 417

1735 AGACCGGAGCCCATTAAGCAGCATGAAC 1765
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
417 rGluArgSerSerHisLysAspSerIleAsn 427

```

seq_name: pir2:A40440

seq_documentation_block:

endothelin 1 and 2 receptor precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 16-Jul-1999

C;Accession: A40440; B49197

R;Lin, H.Y.; Kajl, E.H.; Winkel, G.K.; Ives, H.E.; Lodish, H.F.

Proc. Natl. Acad. Sci. U.S.A. 98, 3185-3189, 1991

A;Title: Cloning and functional expression of a vascular smooth muscle endothelin

A;Reference number: A40440; MUID:91195317

A;Accession: A40440

A;Molecule type: mRNA

A;Residues: 1-426 <LIN>

A;Cross-references: GB:M60786; NID:g204023; PIDN:AAA41114.1; PID:g204024

R;Hori, S.; Komatsu, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.

Endocrinology 130, 1885-1895, 1992

A;Title: Distinct tissue distribution and cellular localization of two messenger r

A;Reference number: A49197; MUID:92191882

A;Accession: B49197

A>Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-45,'N',47-52,'Q',54-426 <HOR>
A:Experimental source: lung
A>Note: sequence extracted from NCBI backbone (NCBIP:89467)
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-426/Product: endothelin 1 and 2 receptor #status predicted <MAT>

alignment_scores:
Quality: 2103.50 Length: 427
Ratio: 5.044 Gaps: 1
Percent Similarity: 97.658 Percent Identity: 92.506

alignment_block:
US-09-931-157-1 x A40440 ..

Align seg 1/1 to: A40440 from: 1 to: 426

```
485 ATGGAAACCCCTTGGCTCAGGACATCCCTTTGGCTGGCACTGGTTGGATG 534
||||| :|||||: 1 MetGlyValLeuCysPheLeuAlaSerPheTrpLeuAlaLeuValGlyG 17
535 TCCTAATCAGTGATAATCCTCGAGAGATACAGCACAAATCTAAGCAATCATG 584
||||| :|||||: 17 yAlaIleAlaAspAsnAlaGluArgTyrSerIleAlaLeuSerHisV 34
585 TGGATGATTCCACCACTTTTCGTGGCAGACAGCTCAGCTTCCTGGTTACC 634
||||| :|||||: 34 algluAspPheThrProPheProGlyThrGluPheAspPheLeuGlyThr 50
635 ACTCATCAACCACTAATTTGGTCTTACCAGCAATGGCTCAATGCACAA 684
||||| :|||||: 51 ThrLeuArgProAsnLeuAlaLeuProSerAsnGlySerMethisG 67
685 CTTATGGCCACAGCAGCTAAATTAATCTCAGCTTTCAAAATACATTAACA 734
||||| :|||||: 67 ytyrCysProGlnGlnThrLysIleThrAlaPheLysTyrIleAsnT 84
735 CTGTGATATCTTGTAATTTTCATGCTGGGAATGGTGGGAATGCAACT 784
||||| :|||||: 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
785 CTGCTCAGGATATTACCAGAACAAATCTATCAGGAATGGCCCAACGC 834
||||| :|||||: 101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
835 GCTGATAGCCAGTCTTGGCCCTTGGAGACCTTATCTATGTTGGTCATTGATC 884
||||| :|||||: 117 alLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
885 TCCTATCAATGATTTAAGCTGCTGGCTGGCGCTGGCTTTTGATCAC 934
||||| :|||||: 134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHis 150
935 AATGACTTTGGCGTATTTCTTTCAGAGCTGTTCCCTTTTTCAGAGATC 984
||||| :|||||: 151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSe 167
985 CTCGCTGGGATCACCGTCTCAACCTCTCGCTCTTACTGTTGACAGGT 1034
||||| :|||||: 167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
1035 ACAGACAGCTTCCTCCTCGAGCTGCTGGTTCAGGGAATGGGATTCCTTG 1084
||||| :|||||: 184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
1085 GTAACCTGCCATGAAATTTGCTCCATCTGACCTCTCTTATCTCTGCC 1134
||||| :|||||: 201 IleThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
```

```
1135 CATTCCTGAAGCGATTGGCTTCGTCATGGTACCCCTTTGAATATAGGGGTG 1184
||||| :|||||: 217 alLeuProGluAlaIleGlyPheValMetValProPheGluTyrLysGlyG 234
1185 ACAGCATAAACCTGCTATGCTCAATGCCACATCAAAATTCATGAGAGTTC 1234
||||| :|||||: 234 luGlnHisArgThrCysMetLeuAsnAlaThrThrLysPheMetGluPhe 250
1235 TACCAGATGTAAGGACTGGTGGCTCTTCGGGTTCTATTCTCTGATGCC 1284
||||| :|||||: 251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPr 267
1285 CTTGTGTGCTACTGCGATCTTCTACCCCTCATGACTTTGTGAGATGTTGA 1334
||||| :|||||: 267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
1335 ACAGAAGGAATGGCAGCTTGAGAAATTCCTCAGTGAACATCTTAAGCAG 1384
||||| :|||||: 284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
1385 CPTCGAGAAGTGGCAAAAACAGTTTTCTGCTGGTTGTAATTTTGTCTCT 1434
||||| :|||||: 301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
1435 TTGCTGGTTCCCTCTTCACCTTAAGCGTATATTGAAGAAACTGTGTATA 1484
||||| :|||||: 317 ucystTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 ACGAATGGACAAGAACCGATGTGAATTACTTACTTACTTACTTACTGCTCATG 1534
||||| :|||||: 334 spGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350
1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAAACCCCAT 1584
||||| :|||||: 351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAATTTGTTCCAGTCATGCC 1634
||||| :|||||: 367 eaLeuLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TCTGCTGCTGCTTTACCAGTCCAAAGTCTGATGACCTCGTCCCATG 1684
||||| :|||||: 384 euCysCysCysCysHisGlnSerLysSerLeuMetThrSerValProMet 400
1685 AACGGAACAAGCATCCAGTGGAAAGAACCCAGCATCAAAACCAACCAACAC 1734
||||| :|||||: 401 AsnGlyThrSerIleGlnTrpLysAsnGlnGln...AsnHisAsnTh 416
1735 AGACCGGAGCGCCATTAAGGACAGCATGAAC 1765
416 rGluArgSerSerHisLysAspSerMetAsn 426
```

seq_name: pir2.A54126

seq_documentation_block:

endothelin receptor-AX - African clawed frog

C:Species: Xenopus laevis (African clawed frog)

C>Date: 07-Jul-1995 #sequence_revision 07-Jul-1995 #text_change 16-Jul-1999

C:Accession: A54126

R:Kumar, C.; Mwangi, V.; Nuthulaganti, P.; Wu, H.L.; Pullen, M.; Brun, K.; Aiyar, H.

J. Biol. Chem. 269, 13414-13420, 1994

A:Title: Cloning and characterization of a novel endothelin receptor from Xenopus he

A:Reference number: A54126; MUID:94230448

A:Accession: A54126

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-415 <KUM>

A:Cross-references: GB:U06633; NID:g507893; PIDN:AAA19570.1; PID:g507894

A>Note: authors translated the codon ATG for residue 84 as His

C:Superfamily: endothelin receptor B

alignment_scores:

1300	GA	TTCTTACACCCCTCAGTCTGTGAGATGTTGAACACGAAGAATGGCA	1349
261	Y	ValPheTyrThrMetMetThrCysGluMetLeuHisInArgIlyslglsy	278
1350	GC	TTCAGAAATTCCTCAGTCAACATCTTAAGCAGCGCTCAGAGAAGTGGCA	1399
278	e	LeuArgIleAlaLeuSerGluHisLeuIysGlnAArgGluValAla	294
1400	AA	ACAGATTTCTGCTGTGGTTGTAATTTTGGCTCTTTGCTGGTTCCTCT	1449
295	L	ysThrValPheCysLeuValValIlePheAlaLeuCysTrpPheProLe	311
1450	TC	ACTTAGCCGCTATATTGAGAAACATGCTATAACGAATAATGCACAGA	1499
311	u	HisLeuSerArgIleIleIysAsnThrIleTyrAsnGluLeuAspPro	328
1500	AC	CAGTGTGAATTAAGTTTCTTACTGTGCTCATGTGATACATCGGTATT	1549
328	e	rArgCysGluLeuLeuSerPheLeuLeuValMetAspPheIleSerIle	344
1550	AA	CTTGGCAACCATGATTCATGTATTAACCCCATAGCTCTGTATTTTGT	1599
345	A	snLeuAlaAlaLeuAsnSerCysIleAsnProIleAlaLeuTyrPheVa	361
1600	G	AGCAGAAATTTAAAAATTTCTTCCAGTCATGCTCTGCTGGTGCNGTT	1649
361	I	serIysIysPheIysAsnCysPheGlnSerCysLeuCysCysCys	377
1650	AC	CAGTCCAAAGTGTGATGACCTCGGTGCCCATGAACGAACAAAGCATC	1699
378	..	GlnSerLysThrHisIleAsnThrAlaProMetAsnValThrSerIle	393
1700	CA	GCTGGAGAACCCAGTCATAAACACCAACACACACCGGACGACAGCA	1749
394	G	lnTrpLysAsnHisAspGlnAsnTyrTyrGlyAlaAspArgSerIleHi	410
1750	TA	AGCACAGCATGAAC	1765
410	s	lysAspSerIleAsn	415
<p>_name: pir2:SI3425</p> <p>_documentation_block:</p> <p>othelin receptor ETB - rat</p> <p>Species: Rattus norvegicus (Norway rat)</p> <p>Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change</p> <p>Accession: SI3425; A49197</p> <p>Author: Akurai, T.; Yanagisawa, M.; Takuma, Y.; Miyazaki, H.; Kimura, Y.</p> <p>Journal: J Biol Chem 267:348, 732-735, 1990</p> <p>Title: Cloning of a cDNA encoding a non-isopeptide-selective</p> <p>Reference number: SI3425; MUID:91080924</p> <p>Accession: SI3425</p> <p>Feature type: mRNA</p> <p>Residues: 1-441 <SA></p> <p>Cross-references: GB:X57764; NID:g56122; PIDN:CAA40916.1; PID:</p> <p>ori, S.; Komatsu, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.</p> <p>ocinology 130, 1885-1895, 1992</p> <p>Article: Distinct tissue distribution and cellular localization</p> <p>Reference number: A49197; MUID:92191882</p> <p>Accession: A49197</p> <p>Status: preliminary; not compared with conceptual translation</p> <p>Molecule type: nucleic acid</p> <p>Residues: 1-65, 'SSAP', 69-441 <HOR></p> <p>Experimental source: lung</p> <p>Source: sequence extracted from NCBI backbone (NCBI:89468)</p> <p>Superfamily: endothelin receptor B</p> <p>Keywords: G protein-coupled receptor; transmembrane protein</p>			
<p>gment_scores:</p> <p>Quality: 1240.00 Length: 359</p> <p>Ratio: 3.949 Gaps: 5</p>			

```
alignment_scores:
  Quality: 1240.00      Length: 359
  Ratio: 3.949          Gaps: 5
  Percent Similarity: 87.465  Percent Identity: 62.674
```


A: Experimental source: lung
A: Note: Sequence extracted from NCBI backbone (NCBIN:125754, NCBIPI:125755)
C: Comment: This endothelin receptor is non-isopeptide selective.

C: Genetics:

A: Gene: GDB:EDNRB; ETB

A: Cross-references: GDB:129075; OMIM:131244

A: Map position: 13q22-13q22

C: Superfamily: endothelin receptor B

C: Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane protein

F: 1-26/Domain: signal sequence #status predicted <SIG>

F: 27-42/Domain: endothelin receptor B #status predicted <NAT>

F: 102-127/Domain: transmembrane #status predicted <TM1>

F: 138-163/Domain: transmembrane #status predicted <TM2>

F: 176-197/Domain: transmembrane #status predicted <TM3>

F: 219-243/Domain: transmembrane #status predicted <TM4>

F: 272-296/Domain: transmembrane #status predicted <TM5>

F: 325-350/Domain: transmembrane #status predicted <TM6>

F: 363-389/Domain: transmembrane #status predicted <TM7>

F: 58/Binding site: carbohydrate (Asn) (covalent) #status predicted

F: 305,413,419,436,440/Binding site: phosphate (Ser) (covalent) #status predicted

alignment_scores:

Quality: 1240.00 Length: 359

Ratio: 3.987 Gaps: 5

Percent Similarity: 86.630 Percent Identity: 63.510

alignment_block:

US-09-931-157-1 x JQ1042 ..

Align seg 1/1 to: JQ1042 from: 1 to: 442

689 TGCCACGACGACTAAATCTTACGCTTTCAATATACATTAACACTGT 738

90 CysGlnGlyProIleGluIleLysGluThrPheLysTyrIleAsnThrVa 106

739 GATATCTTGACTATTTCATCGCTGGGAATGGTGGGAATGCACTCTGC 788

106 IValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeu 123

789 TCAGATCATTTACCAGAACAAATGATGAGGAATGGCCCAACCGCTG 838

123 euArgIleIleTyrIleAsnLysCysMetArgAsnGlyProAsnIleLeu 139

839 ATAGCCAGCTTGCCCTTGAGACCTTATCTATGTGTCATGATCTCC 888

140 IleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleValIleAspIlePr 156

889 TATCAATGATTTAAGCTGCTGGCTGGCGCTGTCCTTTTGTATCACAATG 938

156 OIleAsnValTyrLysLeuLeuAlaGluAspTrpPro..... 168

939 ACTTTGGCGTATTTCTTTGCAAGCTGTTCCTTTTGCAGAACTCTCG 988

169 ..PheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAlaSer 184

989 GTGGGATCAGCTCTCAACCTCTGCGCTCTTACTGTGTGAGGTACAG 1038

185 ValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrAr 201

1039 ACCAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1088

201 gAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrp 218

1089 CTGCCATGAATTCCTCCATCGATCGCTGCTGCTGCTGCTGCTGCTGCT 1138

218 hrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVal 234

1139 CTTGAAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1188

235 ProGluAlaIleGlyPheAspIleIleThrMetAspTyrLysGlySerTy 251

1189 GCATAAACCTGTATGCTCAATGCGCACATCAAAA.....TTCATGGAGT 1232

251 rLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetGlnP 268

1233 TCTACCAAGATGTAAGGACTGGTGGCTCTTCGGGCTTCTATTCTGTATG 1282

268 heTyrLysThrAlaLysAspTrpLeuPheSerPheTyrPheCysLeu 284

1283 CCCTGGTGTGCATCGCATCTTACACCTCATGCTGTGTGAGATGTT 1332

285 ProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGluMetLe 301

1333 GAACAGAAGGAATGCGACCTTGAGAAATGCGCTCAGTCAGTGAACATTAAGC 1382

301 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 317

1383 AGCGTCGAGAAGTGCACAAACAGTTTCTGCTGCTGCTGCTGCTGCTGCT 1432

317 InArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 333

1433 CTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1482

334 LeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTy 350

1483 TAACGAAATGCGACAGAACCGATGCTGAATTAATTTAGTTTCTTACTGCTCA 1532

350 rAsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuVal 367

1533 TGGATTACATCGGTATTAACTTGGCAACCATGATTAATTCATGATAAACCC 1582

367 euAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 383

1583 ATAGCTCTGTATTTGTGAGCAAGAAATTTAAATTTGTTCCATCATG 1632

384 IleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCy 400

1633 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1682

400 sLeuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGluG 416

1683 TGAACGGAACAGCATCCAGTGGAAAGAACCCAGCATCAAAACACCAAC 1732

416 luLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrAsp 431

1733 ACAGACCGGACGACCCATTAAGACAGC 1759

432 AsnPheArgSerSerAsnLysTyrSer 440

seq_name: pir2:157950

seq_documentation_block:

nonselective-type endothelin receptor - rat

C: Species: Rattus sp. (rat)

C: Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999

C: Accession: 157950

R: Cheng, H.F.; Su, Y.M.; Yeh, J.R.; Chang, K.J.

Mol. Pharmacol. 44, 533-538, 1993

A: Title: Alternative transcript of the nonselective-type endothelin receptor from

A: Reference number: 157950; MUID: 93382424

A: Accession: 157950

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: mRNA

A: Residues: 1-442 <RES>

A: Cross-references: GB:S65355; NID:g410692; PIDN: AAB28172.1; PID: g410693

C: Superfamily: endothelin receptor B

alignment_scores:

Quality: 1240.00 Length: 359

Ratio: 3.949 Gaps: 5

Percent Similarity: 87.465 Percent Identity: 62.674

alignment_block:

US-09-931-157-1 x 157950 ..

Align seg 1/1 to: 157950 from: 1 to: 442

```
689 TGCCACGACGACACTAAATTTACTTCAGCTTTCACAAATACATTAACACTGT 738
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
90 CysGlnArgLysIleGluLeuAsnLysThrPheLysThrIleAsnThrIle 106
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
739 GATATCTTCTACTATTTTCATCGTGGGAATGGTGGGAATGCAACACTGC 788
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
106 eValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeu 123
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
789 TCAGGATCATTTACGAGAACAAATGATGAGGAATGGCCCCAACGCGTG 838
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
123 euArgIleIleThrLysAsnLysCysMetArgAsnGlyProAsnIleLeu 139
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
839 ATAGCCAGCTTTGCCCTTGGAGACCTATCTATGTGGTCATTCATCTCC 888
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
140 IleAlaSerLeuAlaLeuGlyAspLeuHisIleIleAspIlePr 156
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
889 TATCAATGTATTTAAGCTGCTGGCTGGCGCTGGCTTTTGTATCAATG 938
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
156 IleAsnAlaThrLysLeuLeuAlaGlyAspTrpPro..... 168
939 ACTTTGGCGTATTTCTTTCAGAGCTGTTCCCTTTTTCGAGAGTCTCTCG 988
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
169 ..PheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAlaSer 184
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
989 GTGGGATCACCGTCTCAACCTCGCGCTCTAGTGTTCAGAGTACAG 1038
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 ValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTrp 201
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1039 AGCAGTGTGCTCTCTGAGTCGTGTTCCAGGGAATGGGATTCCTTTGTTAA 1088
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 galValAlaLaserTrpSerArgIleLysGlyIleGlyValProLysTrp 218
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1089 CTGCCATTTGAATGTCCTCCATCGATCGCTGCTTCCTTTATCTGCGCAT 1138
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 hrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVal 234
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1139 CCGTAGCGCATTTGGCTTCGTCATCGTACCTTGTATATAGGGTCAACA 1188
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
235 ProGluAlaIleGlyPheAspValIleThrSerAspThrLysGlyLysPr 251
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1189 GCATAAAGCTGTATCTCAATGCCACATCAAAA.....TTCATGAGT 1232
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
251 OleuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMetGlnP 268
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1233 TCTACCAAGATGTAAGGACTGTGGCTCTTCGGGTCTATTTCCTGTATG 1282
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 heThrLysThrAlaLysAspTrpTrpLeuPheSerPheThrCysLeu 284
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1283 CCTTGGTGTGCGATCTCTACACCCCTCATCGTGTGAGATGTT 1332
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
285 ProLeuAlaIleThrAlaIlePheThrLeuMetThrCysGluMetLe 301
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1333 GAACGAAGAATGGCAGCTTGAGATTCGCTCAGTGAACATCTTAAGC 1382
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 317
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1383 AGCGTGGAGAAGTGGCAAAACAGTTTCTCGCTGTTGTTGTAATTTTGTCT 1432
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
317 InArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 333
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1433 CTTTGTGCTTCCCTCTTCACTTAACCGGTATATTGAAGAAACTGTGTA 1482
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 LeuCysTrpProLeuHisLeuSerArgIleLeuLysLeuThrLeuTy 350
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1483 TRACGAATGGCAAGACCGATGTGAATTAATTAATTAATTAATTAATTA 1532
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
350 rAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuVal 367
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1533 TGGATTACATCGGTATTAATTAATTAATTAATTAATTAATTAATTAAT 1582
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
367 euAspThrIleGlyIleAsnMetalSerLeuAsnSerCysIleAsnPro 383
1583 ATAGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTTCCAGTCATG 1632
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 IleAlaLeuThrIleuValSerLysArgPheLysAsnCysPheLysSerCy 400
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1633 CCTCTGCTGCTGCTGTACCAAGTCCAAAGTCTGTGATGACCTCGTCCCA 1682
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
400 sLeuCysCysTrpCys...GlnThrPheGluGluLysGlnSerLeuGluG 416
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1683 TGRACCGAACAAGCATCCAGTGGAGAACCAACCATCAAAACAACCAAC 1732
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
416 LuLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrAsp 431
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
1733 ACAGACCGGAGCAGCATTAAGGACAGC 1759
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
432 AsnPheArgSerSerAsnLysTyrSer 440
||| : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

seq_name: pir2:A41591

seq_documentation_block:

C:Species: Bos primigenius taurus (cattle)

C:Date: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change 20-Jun-2000

C:Accession: A41591; A61251; A41184; S63513

R:Saito, Y.; Mizuno, T.; Itakura, M.; Suzuki, Y.; Ito, T.; Hagiwara, H.; Hirose, S.

J. Biol. Chem. 266, 23433-23437, 1991

A:Title: Primary structure of bovine endothelin ET-B receptor and identification of

A:Reference number: A41591; MUID:92078223

A:Accession: A41591

A:Molecule type: mRNA

A:Residues: 1-441 <S>

A:Cross-references: GB:D90456; NID:q217553; PIDN:BAAL4422.1; PID:q217554

R:Hagiwara, H.; Kozuka, M.; Sakaguchi, H.; Eguchi, S.; Ito, T.; Hirose, S.

J. Cardiovasc. Pharmacol. 17(Suppl.7), S117-S118, 1991

A:Title: Separation and purification of 34- and 52-kDa species of bovine lung endoth

A:Reference number: A61251; MUID:92219662

A:Accession: A61251

A:Molecule type: protein

A:Residues: 304-315;424-432 <HAG>

R:Kozuka, M.; Ito, T.; Hirose, S.; Lodhi, K.M.; Hagiwara, H.

J. Biol. Chem. 266, 16892-16896, 1991

A:Title: Purification and characterization of bovine lung endothelin receptor.

A:Reference number: A41184; MUID:91358493

A:Accession: A41184

A:Molecule type: protein

A:Residues: 124-127;S',263-269;304-316;418-421;424-432 <KOZ>

R:Hick, S.; Heidemann, I.; Soskic, V.; Mueller-Esterl, W.; Godovac-Zimmermann, J.

Eur. J. Biochem. 234, 251-257, 1995

A:Title: Isolation of the endothelin B receptor from bovine lung. Structure, signal

A:Reference number: S63513; MUID:96096746

A:Accession: S63513

A:Molecule type: protein

A:Residues: 1, X',3-7, X',9-10;25-35;36-42;73-78;79-87;284-290;291-296, X',298 <HIC>

C:Superfamily: endothelin receptor B

C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:1-36/Domain: signal sequence #status experimental <SIG>

F:27-441/Product: endothelin receptor B #status experimental <MAT>

F:101-126/Domain: transmembrane #status predicted <TM1>

F:137-162/Domain: transmembrane #status predicted <TM2>

F:175-196/Domain: transmembrane #status predicted <TM3>

F:218-242/Domain: transmembrane #status predicted <TM4>

F:271-295/Domain: transmembrane #status predicted <TM5>

F:324-349/Domain: transmembrane #status predicted <TM6>

F:362-388/Domain: transmembrane #status predicted <TM7>

F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:

Quality: 1221.50

Ratio: 3.866

Percent Similarity: 84.043

Length: 376

Gaps: 6

Percent Identity: 60.106

alignment_block:

US-09-931-157-1 x A41591 ..

Align seg 1/1 to: A41591 from: 1 to: 441

```
659 CTACCCAGCAATGCTCAATGCACAACTAT..... 688
72 IleProArgGlyGlyArgMetAlaGlyIleProProArgThrProProPr 88
689 .TGCCACAGCAGACTAAATTTACTTTCAGCTTTCAAAATACATTAACACTG 737
88 oCysaspGlyProIleGluIleLysGluThrPheLysThrIleAsnThrV 105
738 TGATATCTTGACATTTTTCATCTGGAATGGTGGGAATGCAACTCTG 787
105 alValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeu 121
788 CTCAGGATCATTTACCAAGAACAAATGTATAGGAATGCCCAACGCGCT 837
122 LeuArgIleIleTyrlLysAsnLysCysMetArgAsnGlyProAsnIleLe 138
838 GATAGCCAGTCTGCCCTTGGAGACCTTATCTATGTGTCTATGATCTCC 887
138 uIleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleIleIleAspIleP 155
888 CTATCAATGTATTTAAGCTGCTGGCTGGCGCTGGCCTTTTGTATCACAA 937
155 rolleAsnThrTyrlLysLeuLeuAlaLysAspTrpPro..... 167
938 GACTTTGGCGATTCTTTTGAAGCTGTTCCTCTTTTTCAGAAAGTCTCT 987
168 ...PheGluValGluMetCysLysLeuValProPheIleGlnLysAlaSe 183
988 GGTGGGGATCACCGCTCTCAACCTCTGCGCTCTTAGTGTTCACAGGTACA 1037
183 rValGlyIleThrValLeuSerLeuLysAlaLeuSerIleAspArgTyra 200
1038 GAGCAGTTGCTCTGAGTCTGTTCAGGAATGGATTCCTTTGGTA 1087
200 rGAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrp 216
1088 ACTGCCATTGAATGTCTCCTCATGCTGATCTGCTCTTTATCTCTGGCCAT 1137
217 ThrAlaValGluIleValLeuIleTrpValSerValLeuAlaVala 233
1138 TCCTGAAGCGATTTGCTTCATGCTACCTTTGAATATAGGGGTGAAC 1187
233 lProGluAlaValGlyPheAspIleIleThrSerAspHisIleGlyAsnL 250
1188 AGCATAAACCTGTATGCTCAATGCCACATCAAAA.....TTCATGAG 1231
250 ysLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMetGln 266
1232 TTCTACCAAGATGTAAAGACTGGTGGCTCTCGGGTCTTATTTCTGTAT 1281
267 PheTyrlLysThrAlaLysAspTrpIlePheSerPheTyrlPheCysLe 283
1282 GCCCTTGCTGCTGCTGCTGCTCTACACCTCTCATGCTTGTGAGATGT 1331
283 uProLeuAlaIleThrAlaLeuPheTyrlThrLeuMetThrCysGluMetL 300
1332 TGAACAGAGGAATGGCAGCTTGAGAATGGCCCTCAGTGAACATCTTAAG 1381
300 euArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLys 315
1382 GAGCCTCAGAGTGGCAAAACAGCTTTCTGCTTGTGTGTAATTTTTC 1431
316 GlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAl 332
1432 TCTTTGCTGGTCTCTCTTCACTTAAGCGGTATATTGAAGAAACCTGTG 1481
332 aleuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuT 349
```

```
1482 ATACGAAATGCACAGACCGATGTGAATTACTTAGTTCTTCTTACTGTC 1531
349 yrAspGlnHisAspProArgArgCysGluPheLeuSerPheLeuVal 365
1532 ATGATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAAACC 1581
366 LeuAspTyrlleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPr 382
1582 CATAGCTCTGTATTTTGTGACAGAAATTTAAATTTTCCAGTCAT 1631
382 oIleAlaLeuTyrlLeuValSerLysArgPheLysAsnCysPheLysSerC 399
1632 GCCTCTGCTGCTGTATCCAGTCCCAAAAGCTCTGATGACCTCGTCCCC 1681
399 ysLeuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGlu 414
1682 ATGAACGACAAACATCCAGTGAAGAACCCAGCATCAAAACACACCAA 1731
415 GlulysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyraS 430
1732 CACAGCGGAGCAGCCATTAAGGACAGC 1759
430 pAsnPheArgSerSerAsnLysTyrlSer 439
seq_name: pir2:A48538
```

seq_documentation_block:
endothelin-3 receptor precursor - African clawed frog
C:Species: Xenopus laevis (African clawed frog)
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 16-Jul-1999
C:Accession: A48538
R:Karne, S.; Jayawickreme, C.K.; Lerner, M.R.
J. Biol. Chem. 268, 19126-19133, 1993
A:Title: Cloning and characterization of an endothelin-3 specific receptor (ET-3 r
A:Reference number: A48538; MUID:93366838
A:Accession: A48538
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-444 <RAR>
A:Cross-references: GB:L20299; NID:9402938; PIDN:AAA9704.1; PID:9402939
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; transmembrane protein

alignment_scores:

Quality: 1014.50 Length: 392
Ratio: 3.404 Gaps: 11
Percent Similarity: 76.020 Percent Identity: 52.041

alignment_block:

US-09-931-157-1 x A48538 ..

Align seg 1/1 to: A48538 from: 1 to: 444

```
665 AGCAATGCTCAATGCACAACTAT.....TGCCC 693
62 SerThrGlyAsnValLeuAsnMetSerProProSerProCysLe 78
694 ACAGCAGACTAAATTTACTTCCAGCTTTCAAAATACATTAACACTGTGATAT 743
78 uSerArgAlaLysIleArgHisAlaPheLysTyrlValThrThrIleLeuS 95
744 CTTGTACTATTTTTCATCTGGAATGGTGGGAATGCAACTCTGCTCAGG 793
95 erCysValIlePheLeuValGlyIleValGlyAsnSerThrLeuLeuArg 111
794 ATCATTTTACCAGAACAAATGTATAGGAATGCCCAACGCGCTCATAGC 843
112 IleIleTyrlLysAsnLysCysMetArgAsnGlyProAsnValLeuIleAl 128
844 CAGTCTTCCCTTGGAGACCTTATCTATGTGTGCTCATGATCTCCCTATC 892
128 aserLeuAlaLeuGlyAspLeuPheTyrlleLeuIleAlaIleProIleI 145
```

```
893 .....AATGTAATTAGCTGGCTGGCGCTGGCTTTTGATCACAAT 937
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
145 leSerIleSerPheTrpLeuSerThrGly.....HisSer 156
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
938 GACTTTGGCGTATTTCTTTGGCAAGCTGTTCCCTTTTTCGACGAAGTCCTC 987
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
157 GluTyr.....IleTyrGlnLeuValHisLeuTyrAr 167
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
988 GGTGGGGATCACCGTCTCAACCTCTGCGCTCTTAGTGTTCGACAGTACA 1037
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
167 gAlaArgValTyrSerLeuSerLeuCysAlaLeuSerIleAspArgTyrA 184
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1038 GAGCAGTTGCCCTCCCTGGAGTCGTGTTCAGGAAATGGGATTCCTTTGGTA 1087
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
184 gAlaValAlaSerTrpAsnArgIleArgSerIleGlyIleProValArg 200
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1088 ACTGCCATTGAATGTCCTCCATCTGGATCCCTGTCCTTTATCTGGCCAT 1137
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
201 LysAlaIleGluLeuThrLeuIleTrpAlaValAlaIleIleValAlaVa 217
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1138 TCCTGAAGCGATTGGCTTCGTATGTCATGTCACCTTGAATATAGGGTGAAC 1187
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
217 IProGluAlaIleAlaPheAsnLeuValGluLeuAspPheArgGlyGlnT 234
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1188 AGCATAAAACCTGTATGCTC.....AATGCCACATCAAAATCATGGAG 1231
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
234 hrIleLeuValCysMetLeuProMetGluGlnThrSerAspPheMetArg 250
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1232 TCTTACCAGATGTAAGACGCTGGTCTTCCTGGGTTCTATTCTGTAT 1281
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
251 PheTyrGlnGluValLysValTrpTrpLeuPheGlyPheTyrPheCysLe 267
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1282 GCCCTGTGTGTCACGCTGCTTCTACACCTCATGACTTGTGAGATGT 1331
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
267 aProLeuAlaCysThrGlyValPheTyrThrLeuMetSerCysGluMetL 284
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1332 TGAACAGGAAGTAATGGCAGCTTGAGAATTCGCCCTCAGTGAACATCTTAAG 1381
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
284 euSerIleLysAsnGly...MetArgIleAlaLeuAsnAspHisMetLys 299
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1382 CAGCTCGAGAAAGTGGCAAAACAGTTTCTGCTGGTGTGTAATTTTGC 1431
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
300 GlnArgArgGluValAlaLysThrValPheCysLeuValValIlePheAl 316
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1432 TCTTTGCTGTTCCCTCTTCACTTAAGCGGTATATG..... 1468
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
316 aLeuCysTrpLeuProLeuHisValSerSerIlePheValArgLeuSera 333
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1469 .....AAGAAACTGTGTATTAAC 1486
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
333 laThrValLysArgAlaCysIleLeuLysAsnLysArgSer..CysIleMe 349
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1487 GAAATGGACAAGACCGATGTGAATTACTTACTTCTTACTGCTCATGGA 1536
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
349 talacIleIleGlnThrGlyValAsnTyr..GlnLeuLeuMetValMetAs 365
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1537 TTACATCGGTATTACTTGGCAACCATGAATTCATGTATAAACCCCATAG 1586
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
365 nTyrThrGlyIleAsnMetAlaSerLeuAsnSerCysIleGlyProValA 382
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1587 CTCTCTATTTCGACGAAGAATTAATAATTTTCCAGTCATCCCTC 1636
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 laLeuTyrPheValSerArgLysPheLysAsnCysPheGlnSerCysLeu 398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1637 TGCTGCTGCTGTACAGTCACAAAAGTCTGATGACCTGGTCCCATGAA 1686
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
399 CysCysTrpCysHisArgProThrLeuThrIleThrProMetAspGluLy 415
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1687 CGGAACAAGCATCTCAGTGAAGAACCCAGATCAAAACACACAC.....A 1730
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
415 sGlySerGlyGlyLysTrpLys.....AlaAsnGlyHisAspLeuA 429
```

```
1731 ACACAGACCGGAGCAGCCATAAG 1753
      :: ||||| ||||| ::
429 spLeuAspArgSerSerArg 436
seq_name: pir2:B57508
seq_documentation_block:
endothelin-1 receptor homolog - mouse (fragment)
N:Alternate names: G-protein coupled receptor Gpcr10
C:Species: Mus musculus (house mouse)
C>Date: 01-Dec-1995 #sequence_revision 01-Dec-1995 #text_change 16-Jul-1999
R:Wilkie, T.M.; Chen, Y.; Gilbert, D.J.; Moore, K.J.; Yu, L.; Simon, M.I.; Copeland,
Genomics 18, 175-184, 1993
A:Title: Identification, chromosomal location, and genome organization of mammalian
A:Reference number: A48909; MUID:94116980
A:Accession: B57508
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-117 <WIL>
A:Cross-references: GB:L20340; NID:q438784; PIDN:AAA16845.1; PID:q438785
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor
```

```
alignment_scores:
Quality: 605.00 Length: 117
Ratio: 5.216 Gaps: 0
Percent Similarity: 99.145 Percent Identity: 96.581
alignment_block:
US-09-931-157-1 x B57508 ..
Align seg 1/1 to: B57508 from: 1 to: 117
1076 ATTCTTTGGTAAGTCCCATTTGAAATTTGCTCATCTGGATCTGCTGCTT 1125
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1 IleProLeuIleThrAlaIleGluIleValSerIleTrpIleLeuSerPh 17
1126 TATCTGGCCATTCCTGAAGCGATTGGCTGCTCATGCTGCTGCTTGAAT 1175
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
17 eIleLeuAlaIleProGluAlaIleGlyPheValMetValProPheGlu 34
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1176 ATAGGGGTGAACAGCATAAAACCTGTATGCTCAATGCCACATCAAAATTC 1225
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
34 YrLysGlyGluLeuHisArgThrCysMetLeuAsnAlaThrSerLysPhe 50
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1226 ATGGAGTTCTACCAAGATGTAAGGACTGGTGGCTCTTCGGGTTCTATTT 1275
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
51 MetGluPheTyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPh 67
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1276 CTGTATGCCCTTGGTGTGCTGCTGCGATCTTCTACACCTCATGCTGTG 1325
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
67 eCysMetProLeuValCysThrAlaIlePheTyrThrLeuMetThrCysG 84
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1326 ACATGTTGAACAGGAAGTGGCAGCTTGAGAAATTCCTCCTCAGTGAACAT 1375
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
84 LuMetLeuAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHis 100
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1376 CTTAAGCAGCGTCGAGAAGTGGCAAAACAGTTTCTGCTTGGTTGTTAAT 1425
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
101 LeuLysGlnArgArgGluValAlaLysThrValPheCysLeuValIle 117
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1426 T 1426
      |
117 e 117
```

```
seq_name: pir2:A39003
seq_documentation_block:
bombesin/gastrin-releasing peptide receptor [similarity] - mouse
C:Species: Mus musculus (house mouse)
C>Date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change 20-Apr-2000
```

C:Accession: A39003; A36553
R:Battey, J.F.; Way, J.M.; Corjay, M.H.; Shapira, H.; Kusano, K.; Harkins, R.; Wu, J.M.;
Proc. Natl. Acad. Sci. U.S.A. 88, 395-399, 1991
A:Title: Molecular cloning of the bombesin/gastrin-releasing peptide receptor from Swiss
A:Reference number: A39003; MUID:91110536
A:Accession: A39003
A:Molecule type: mRNA
A:Residues: 1-384 <BAT>
A:Cross-references: GB:M57922; NID:g996089; PIDN:AA75650.1; PID:g996090
A:Note: the sequence in GenBank entry MUSGRPBM, release 113.0, has the codon CCG for 30
Mol. Endocrinol. 4, 1956-1963, 1990
A:Title: Cloning and functional characterization of a complementary DNA encoding the mur
A:Reference number: A36553; MUID:91187004
A:Accession: A36553
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-308, 'P', 310-384 <SPT>
A:Cross-references: GB:M61000; GB:M35328; NID:g193650; PIDN:AAA37744.1; PID:g193651
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transme
F:40-63/Domain: transmembrane #status predicted <TR1>
F:78-97/Domain: transmembrane #status predicted <TR2>
F:116-137/Domain: transmembrane #status predicted <TR3>
F:154-174/Domain: transmembrane #status predicted <TR4>
F:210-235/Domain: transmembrane #status predicted <TR5>
F:267-287/Domain: transmembrane #status predicted <TR6>
F:299-326/Domain: transmembrane #status predicted <TR7>
F:20,24,59,191,347/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:
Quality: 465.00 Length: 342
Ratio: 2.076 Gaps: 7
Percent Similarity: 65.497 Percent Identity: 32.456

alignment_block:
US-09-931-157-1 x A39003 ..

Align seg 1/1 to: A39003 from: 1 to: 384

719 TTCAAATACATTAAACACGTGATATCTGTACTATTTTCATCGTGGGAAT 768
40 PheIleTyValIleProAlaValTyGlyLeuIleIleValIleGlyLe 56
769 GGTGGGGAATGCACTCTGCTCAGGATCATTTACCAGAACAAATGATGA 818
56 uIleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMetA 73
819 GGAATGGCCCAACGGCGTGATAGCAGCTTGGCCTTGGAGACCTTATC 868
73 rgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeu 89
869 TATGTGGTCATTGATCTCCCTATCAATGATTTAAAGCTCGCTGGGCG 918
90 LeuLeuValThrCysAlaProValAspAlaSerLysTyLeuAlaAsp 106
919 CTGGCCTTTTGATCAACAATGACTTTGGCGTATTTCTTTGCAAGCTGTTC 968
106 grpLeu.....PheGlyArgIleGlyCysLysLeuIleP 118
969 CTTTTCGAGAGCTCGGTGGGATACCGCTCTCAACCTCGCGCT 1018
118 roPheIleGlnLeuThrSerValGlyValSerValPheThrLeuThrAla 134
1019 CTTAGTGTTCAGGTACAGACAGTTCCTCTCGAGTCGTTCAGG 1068
135 LeuSerAlaAspArgTyLysAlaIleValArgProMetAspIleGlnAl 151
1069 AATGGGATCTTTGGTAACGCCATTGAAATTTGTCCTCATCTGGATCC 1118
151 aserHisAlaLeuMetLysIleCysLeuLysAlaIleLeuIleTrpIle 168

1119 TGTCTTTATCTCGCCATCTCCTGAAGGATTGCTTCTGTCATGTTACCC 1168
168 alSerMetLeuAlaIleProGluAlaValPheSerAspLeuHisPro 184
1169 TTTCAAATATAGGGTGACACACATAAACCTGTATGCTCAATGCCACATC 1218
185 PheHisValLysAspThrAsnGlnThrPheIleSerCysAlaProTyPr 201
1219 AAAATTCATGAGTTCATACCAAGATGTAAAGACTGGTGGCTCTTCGGGT 1268
201 oHisSerAsnGluLeuHisProLysIleHisSerMetAlaSerPheLeu 218
1269 TCTATTTCTGATGCTTGTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318
218 alPheTyValIleProLeuAlaIleIleSerValTyTyTyPheIle 234
1319 ACTTGTGAGATGTTGAACAGAGG.....AATGGCAGCTT 1353
235 AlaArgAsnLeuIleGlnSerAlaTyArgLeuProValGluGlyAsnI 251
1354 GAGAAATCCCTCAGTGAACATCTTAAACAGCGTCGAGAAGTGGCAAAA 1403
251 eHisVal.....LysLysGlnIleGluSerArgLysArgLeuAlaLys 266
1404 CAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
266 hrValLeuValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHis 282
1454 TTAAGCCCTATATGAAGAAACCTGTGTATTAACGAAATGGACAAACCG 1503
283 ValIleTyLeuTyArgSerTyHisTySerGluValAspThrSer... 298
1504 ATGTGAATTAATCTGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1553
299MetLeuHisPhe.....ValThrSerIleCysAlaHisLeu 311
1554 TGGCAACCATGAATTCATGATATAACCCCATGCTGTGTATTTCTGAGC 1603
311 euAlaPheThrAsnSerCysValAsnProPheAlaLeuTyLeuLeuSer 327
1604 AAGAAATTTAAATTTGTTCCAGTCATGCTGCTGCTGCTGCTGCTGCTGCT 1653
328 LysSerPheArgLysGlnPheAsnThrGlnLeuLeuCysCys..... 341
1654 GTCCAAAGCTGATGACCTCGTCCCTCCCATGACGGAACCAACATCCAGT 1703
342 .GlnProGlyLeuMetAsn.ArgSerHisSerThrGlyArgSerThrThr 357
1704 GGAAGAACCCATCAACAAACACC 1727
358 CysMetThrSerPheLysSerThr 365
seq_name: pir2:157682
seq_documentation_block:
bombesin/ GRP receptor - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999
C:Accession: 157682
R:Wada, E.; Way, J.; Shapira, H.; Battey, J.F.
Mol. Cell. Neurosci. 3, 446-460, 1992
A:Title: Two distinct bombesin receptor subtype subtypes in ppostnatal rat central
A:Reference number: 157682
A:Accession: 157682
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <RES>
A:Cross-references: EMBL:X56661; NID:g509398; PIDN:CAA39988.1; PID:g509399
C:Superfamily: endothelin receptor B

alignment_scores:
Quality: 463.00 Length: 340

Ratio: 2.095 Gaps: 6
Percent Similarity: 65.000 Percent Identity: 31.765
alignment_block:
US-09-931-157-1 x I57682 ..
Align seg 1/1 to: I57682 from: 1 to: 384
725 TACATTAACTGTGTACTATTTCATCTGGGATGTGGG 774
||||| : : : : : ||| : : : : :
42 TyrValIleProAlaValTyrGlyLeuIleValIleGlyLeuIleG1 58
775 GAATGCAACTCTGCTCAGGATCATTTTACCAGAACAAATGTATGAGGAATG 824
||||| : : : : : ||| : : : : :
58 yAsnIleThrLeuIleLysIlePheCysThrValLysSerMetArgAsnV 75
825 GCCCAACCGCTGTAGCAGCTCTTGGCCTGGAGACCTTATCTATGTG 874
||||| : : : : : ||| : : : : :
75 alProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLeuLeu 91
875 GTCAATTGATCTCCCTATCAATGTATTAAAGCTGCTGGCTGGCGCTGGCC 924
||| : : : : : ||| : : : : :
92 ValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspArgTrpLe 108
925 TTTTGATCAATGACTTTGGGCTATTCTTTGCAAGCTGTTCCCTTTT 974
||||| : : : : : ||| : : : : :
108 u.....PheGlyArgIleGlyCysLysLeuIleProPheI 120
975 TGCAGAGTCTCTGGGCGGATCACCCTCAACCTCTGCGCTCTTAGT 1024
||||| : : : : : ||| : : : : :
120 leGlnLeuThrSerValGlyValSerValPheThrLeuThrAlaLeuSer 136
1025 GTTGACAGCTACAGACAGCTGCTCTGAGTCGCTCAGGGAATGG 1074
||||| : : : : : ||| : : : : :
137 AlaAspArgTyrLysAlaIleValArgProMetAspIleGlnAlaSerHI 153
1075 GATTCCTTTGGTAACTGCCATGAATGTCTCCATCTGGATCCCTGCT 1124
||| : : : : : ||| : : : : :
153 sAlaLeuMetLysIleCysLeuLysAlaAlaLeuIleTrpIleValSerM 170
1125 TTATCTGCGCATCTCTGAAGCGATTGGCTGCTCATGTGATCCCTTTGA 1174
||||| : : : : : ||| : : : : :
170 etLeuLeuAlaIleProGluAlaValPheSerAspLeuHisProPheHis 186
1175 TATAGGGGTGAACAGCATAAACCTGTATGCTCAATGCCACATCAAAAT 1224
||| : : : : : ||| : : : : :
187 ValLysAspThrAsnGlnThrPheIleSerCysAlaProTyrProHis 203
1225 CATGGAGTTCTACCAAGATGTAAGGACTGGTGGCTCTTGGGTTCTATT 1274
||||| : : : : : ||| : : : : :
203 rAsnGluLeuHisProLysIleHisSerMetAlaSerPheLeuValPheT 220
1275 TCTGTATGCCCTTGGTGCATCGCATCTTCTACACCTCATGACTTGT 1324
||| : : : : : ||| : : : : :
220 yrIleIleProLeuSerIleIleSerValTyrTyrPheIleAlaArg 236
1325 GAGATCTTGAACAGAAGG.....AATGGCGCTTGAGAT 1359
||| : : : : : ||| : : : : :
237 AsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIleHisVa 253
1360 TGCCCTCAGTGAACATCTTAAAGCAGCGTCGAGAAGTGGCAAAACAGTTT 1409
: : : : : ||| : : : : :
253 l.....LysLysGlnIleGluSerArgLysArgLeuAlaLysThrValL 268
1410 TCTGCTTGGTGTAAATTTTGTCTTGTGGTTCCTCTTCACTTAAGC 1459
||| : : : : : ||| : : : : :
268 euValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIle 284
1460 CGTATATTGAAGAAACTGTGTATAAGCAATGCAAGAACCGATGTGA 1509
||| : : : : : ||| : : : : :
285 TyrLeuTyrArgSerTyrHisTyrSerGluValAspThrSerMetLeuHI 301
1510 ATTACTTAGTTTCTTACTGCTCATGATTACATCGGTATTAACTTGGCAA 1559

301 sPheIleThrSerIleCysAlaArg.....LeuLeuAlap 313
: : : : : ||| : : : : :
1560 CCATGAATTCATCTATATAACCCCATAGCTCTGTATTGTCAGCAAGAA 1609
||||| : : : : : ||| : : : : :
313 heThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSer 329
1610 TTTTAAAAATTTGTTCCAGTCATGCTGCTGCTGCTGCTTACCAGTCCAA 1659
||||| : : : : : ||| : : : : :
330 PheArgLysGlnPheAsnThrGlnLeuLeuCysCys.....GlnPr 343
1660 AACTCTGATGACCTCGTCCCATGACGGAACAAGCATCCAGTGGGAAGA 1709
||||| : : : : : ||| : : : : :
343 oSerLeuLeuAsn.ArgSerHisSerThrGlyArgSerThrThrCysMet 359
1710 ACCACGATCAAAACAACC 1727
360 ThrSerPheLysSerThr 365
seq_name: pir2:A41007
seq_documentation_block:
gastrin-releasing peptide receptor - human
C:Species: Homo sapiens (man)
C:Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 16-Jul-1999
C:Accession: A41007
R:Corjay, M.H.; Dobrzanski, D.J.; Way, J.M.; Viallet, J.; Shapira, H.; Worland, P.;
J. Biol. Chem. 266, 18771-18779, 1991
A:Title: Two distinct bombesin receptor subtypes are expressed and functional in hu
A:Reference number: A41007; MUID:92011639
A:Accession: A41007
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-384 <COR>
A:Cross-references: GB:M73481; NID:g183649; PIDN:AAA88050.1; PID:g183650
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; transmembrane protein
alignment_scores:
Quality: 455.50 Length: 342
Ratio: 2.089 Gaps: 7
Percent Similarity: 63.743 Percent Identity: 30.994
alignment_block:
US-09-931-157-1 x A41007 ..
Align seg 1/1 to: A41007 from: 1 to: 384
725 TACATTAACTGTGTACTATTTCATCTGGGATGTGGG 774
||||| : : : : : ||| : : : : :
41 TyrValIleProAlaValTyrGlyValIleIleLeuIleGlyLeuIleG1 57
775 GAATGCAACTCTGCTCAGGATCATTTTACCAGAACAAATGTATGAGGAATG 824
||||| : : : : : ||| : : : : :
57 yAsnIleThrLeuIleLysIlePheCysThrValLysSerMetArgAsnV 74
825 GCCCAACCGCTGTAGCAGCTCTTGGCCTGGAGACCTTATCTATGTG 874
||||| : : : : : ||| : : : : :
74 alProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLeuLeu 90
875 GTCAATTGATCTCCCTATCAATGTATTAAAGCTGCTGGCTGGCGCTGGCC 924
||| : : : : : ||| : : : : :
91 IleThrCysAlaProValAspAlaSerArgTyrLeuAlaAspArgTrpLe 107
925 TTTTGATCAATGACTTTGGGCTATTCTTTGCAAGCTGTTCCCTTTT 974
||||| : : : : : ||| : : : : :
107 u.....PheGlyArgIleGlyCysLysLeuIleProPheI 119
975 TGCAGAGTCTCTGGTGGGATCACCCTCAACCTCTGCGCTCTTAGT 1024
||||| : : : : : ||| : : : : :
119 leGlnLeuThrSerValGlyValSerValPheThrLeuThrAlaLeuSer 135


```
1025 GTTCACAGGTACAGCAGTGGCTCCTCGAGTGGTTCACGGGAATGG 1074
      ::::::::::::::::::::
136 AlaAspArgTyrLysAlaIleValArgProMetAspIleGlnAlaSerHi 152
      ::::::::::::::::::::
1075 GATTCCTTGGTAACTGCCATGAATGCTCCATCTGATGATCCTGCTCT 1124
      ::::::::::::::::::::
152 sAlaLeuMetLysIleCysLeuLysAlaAlaPheIleTrpIleIleSerM 169
      ::::::::::::::::::::
1125 TTATCTGGCCATCTCTGAAGCAGTGGCTCGTCACTGGTACCCCTTT... 1171
      ::::::::::::::::::::
169 etLeuLeuAlaIleProGluAlaValPheSerAspLeuHisProPheHis 185
      ::::::::::::::::::::
1172 ...GAATATAGGGGTGAACAGCATAAAACCTGTATGCTCAATGCCACATC 1218
      ::::::::::::::::::::
186 GluGluSerThrAsnGlnThrPheIleSerCysAlaProTyrProHisSe 202
      ::::::::::::::::::::
1219 AAAATTCATGAGTCTTACCAAGATGTAAGGACTGCTGCTCTCGGGT 1268
      ::::::::::::::::::::
202 rAsn.....GluLeuHisProLysIleHisSerMetAlaSerPheLeuV 217
      ::::::::::::::::::::
1269 TCTATTCTGTATCGCTTGGTGTGCACCTGCGATCTTCTACACCCCTCATG 1318
      ::::::::::::::::::::
217 alPheTyrValIleProLeuSerIleIleSerValTyrTyrPheIle 233
      ::::::::::::::::::::
1319 ACTTGTGAGATGTTGAACAGAAGG.....AATGCCAGCTT 1353
      ::::::::::::::::::::
234 AlAlaLysAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnI 250
      ::::::::::::::::::::
1354 GAGAAATGCCCTCAGTGAACATCTTAAGCAGCGTCGAGAGTGGCAAAAA 1403
      ::::::::::::::::::::
250 eHisVal.....LysLysGlnIleGluSerArgLysArgLeuAlaLysT 265
      ::::::::::::::::::::
1404 CAGTTTCTGCTGCTGTTGTAATTTTGTCTTCTTGTCTGCTCCCTCTTCAC 1453
      ::::::::::::::::::::
265 hrValLeuValPheValClyLeuPheAlaPheCysTrpLeuProAsnHis 281
      ::::::::::::::::::::
1454 TTAAGCCCGTATATTGAAGAAACCTGTGTATTAACGAATGGACAGAAGCCG 1503
      ::::::::::::::::::::
282 ValIleTyrLeuTyrArgSerTyrHisTyrSerGluValAspThrSerMe 298
      ::::::::::::::::::::
1504 ATGTGAATTACTAGTTCTTACTGCATCATGATGATTACATCGTATTACT 1553
      ::::::::::::::::::::
298 tLeuHisPheValThrSerIleCysAlaArg.....LeuL 310
      ::::::::::::::::::::
1554 TGGCAACCATCATGATGATTAACCCCATAGTCTGTATTGTTGTGAGC 1603
      ::::::::::::::::::::
310 euAlaPheThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSer 326
      ::::::::::::::::::::
1604 AAGAAATTTAAATTTGTTCCAGTCATGCTCTGCTGCTGCTGCTGCTGCT 1645
      ::::::::::::::::::::
327 LysSerPheArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGl 343
      ::::::::::::::::::::
1646 .....TGTTACAGTCCCAAAAGTCTGATGACCT 1673
      ::::::::::::::::::::
343 yLeuIleIleArgSerHisSerThrGlyArgSerThrThrCysMetThrS 360
      ::::::::::::::::::::
1674 CGGTCCCATGACGGAACAGCATC 1699
      ::::::::::::::::::::
360 erLeuLysSerThrAsnProSerVal 368
      ::::::::::::::::::::
```

seq_name: pir2:JH0374

seq_documentation_block:

bombesin receptor, neuromedin-B-preferring - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999

C:Accession: JH0374

R:Wada, E.; Way, J.; Shapira, H.; Kusano, K.; Lebacqz-Verheyden, A.M.; Coy, D.; Jensen, R

Neuron 6, 421-430, 1991

A:Title: cDNA cloning, characterization, and brain region-specific expression of a neuro

A:Reference number: JH0374; M01D:91158987

A:Accession: JH0374

A:Molecule type: mRNA

A:Residues: 1-390 <WAD>
A:Cross-references: GB:U37058; NID:g1039481; PIDN:AAA79881.1; PID:g1039482
A:Experimental source: esophagus
C:Superfamily: endothelin receptor B
F:42-65/Domain: transmembrane #status predicted <TD1>
F:80-99/Domain: transmembrane #status predicted <TD2>
F:118-139/Domain: transmembrane #status predicted <TD3>
F:157-177/Domain: transmembrane #status predicted <TD4>
F:212-235/Domain: transmembrane #status predicted <TD5>
F:267-287/Domain: transmembrane #status predicted <TD6>
F:300-327/Domain: transmembrane #status predicted <TD7>
F:8,71,192/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:
Quality: 454.50 Length: 408
Ratio: 1.959 Gaps: 12
Percent Similarity: 56.863 Percent Identity: 30.392

alignment_block:

US-09-931-157-1 x JH0374 ..

Align seg 1/1 to: JH0374 from: 1 to: 390

```
650 AATTGTGCTCTACCCAGCAATGGCTCA..... 676
      ::::::::::::::::::::
      8 AsnLeuSerLeuProThrGluAlaSerGluSerGluLeuGluProGluVa 24
      ::::::::::::::::::::
677 ....ATCCACAACCTATTGCCACAGCAGCTATAAATTAATCTACGCTTCA 722
      ::::::::::::::::::::
      24 lTrpGluAsnAspPheLeuProAspSerAspGlyThrAla..... 38
      ::::::::::::::::::::
723 AATACATTAACACTGTGATATCTTGTACT..... 751
      ::::::::::::::::::::
      39 .....GluLeuValIleArgCysValIleProSerLeuTyrLeuIle 52
      ::::::::::::::::::::
752 ATTTTCATCTGGGAATGGTGGGAATCCAACCTCTGCTCAGGATCATTTA 801
      ::::::::::::::::::::
      53 IleIleSerValGlyLeuLeuGlyAsnIleMetLeuValLysIlePheLe 69
      ::::::::::::::::::::
802 CCAGACAATATGTATGAGGAATGGCCCAACGCGCTGATAGCCAGCTCTTG 851
      ::::::::::::::::::::
      69 uThrAsnSerThrMetArgSerValProAsnIlePheIleSerAsnLeuA 86
      ::::::::::::::::::::
852 CCCTGGGAGACCTTATCTATGTGTATGATGATGATGATGATGATGATGAT 901
      ::::::::::::::::::::
      86 laAlaGlyAspLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 100
      ::::::::::::::::::::
902 AAGCTGCTGGCTGGCGCTGGCTTTTGCATCACAATGACTTTGGCGTATT 951
      ::::::::::::::::::::
      101 .....AlaSerArgTyrPheAspGluTrpValPheGlyLysLe 114
      ::::::::::::::::::::
952 TCTTTGCAAGCTGTCCCTCTTTTCAGAGAAGTCTCTGGTGGGATCACCG 1001
      ::::::::::::::::::::
      114 uGlyCysLysLeuIleProAlaIleGlnLeuThrSerValGlyValSerV 131
      ::::::::::::::::::::
1002 TCCTCAACCTCTGCGCTCTTATGTTGACAGGTACAGAGCAGCTTCCTCC 1051
      ::::::::::::::::::::
      131 alPheThrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsn 147
      ::::::::::::::::::::
1052 TGGAGTCTGTTACAGGAATGGGATTCCTTTGGTAACGTCATTCGAAT 1101
      ::::::::::::::::::::
      148 ProMetAspMetGlnThrSerGlyValValLeuTrpThrSerLeuLysAl 164
      ::::::::::::::::::::
1102 TGTCTCCATCTGGATCTCTCTTATCTGCGCCATTCCTGAAGCGATT. 1150
      ::::::::::::::::::::
      164 aValGlyIleTrpValValSerValLeuLeuLeuLeuLeuLeuLeuLeuLeu 181
      ::::::::::::::::::::
1151 ..... 1159
      ::::::::::::::::::::
      181 heSerGluValAlaAlaArgIleGlySerSerAspAsnSerSerPheThrAla 197
      ::::::::::::::::::::
```

1160 ATGGTACCTTTTCATATATAGGGCTGAACAGCATATAAACCTGTATGCTCAA 1209
:::|||||
198 CysileProTyrProGlnThrAspGluLeuHis..... 208
1210 TGCCACATCAAAATTCATGGAGTTCTACCAAGATGAAGGACTGGTGGC 1259
209ProLysileHisSerValLeuI 216
1260 TCTTCGGGTTCTATTTCGTATGCTGGTGTGCACCTGCGATCTTCTAC 1309
:::|||||
216 lePheLeuValTyrPheLeuileProLeuValileleSerileTyr 232
1310 ACCCTCATGCTGTGAGATGTTGAACAGCAAGGAATGGAGCTTCAGAA 1359
:::|||||
233 ...TyrHisileAlaLysThrLeuileArgSerAlaHisAsnLeuProI 248
1360 TGCCCTCAGTGAACATCTTAACGAG.....CGTCGAGAAGTGG 1397
248 yGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLysArgLeuA 265
1398 CAAAACAGTTTCTGCTTGGTGTGTAATTTTGTCTTGTGCTGCTCCCT 1447
|||||
265 laLysileValleuValPheValGlyCysPheValPheCysTrpPhePro 281
1448 CTTCACTTAAGCCGTATATTGAAGAAACTGTGTATAACGAATGGACAA 1497
|||||
282 AsnHisileLeuTyrLeuTyrArgSerPheAsnTyrLysGluileAspPr 298
1498 GAACCGATGTAATTACTTAGTTCTTACTGCTCATGATACATCGGTA 1547
:::|||||
298 oSerLeuGlyHisMetileValThrLeu.....ValAlaA 310
1548 TTAACCTGGCAACCATGAATCATATATAACCCCATAGCTCTGTATTTT 1597
|||||
310 rgValLeuSerPheSerAsnSerCysValAsnPropheAlaLeuTyrLeu 326
1598 GTGACGAAGAAATTTAAAAATTTTCCAGTCATGCTCTGCTGC..... 1642
:::|||||
327 LeuSerGluSerPheArgLysHisPheAsnSerGlnLeuCysCysGlyG 343
1643 .TGCTGTACCAGTCCAAAGTCTG..... 1666
|||
343 nLysSerTyrProGluArgSerThrSerTyrLeuLeuSerSerSerAlaV 360
1667ATGACCTCGTCCCTCCATGAACGAACAGCATCCAGTGGGAAGAC 1711
360 alArgMetThrSerLeuLysSerAlaLysAsnValThrAsnSer 376
1712 CACGATCAAAACACCAACACACA 1735
|||
377 ValLeuLeuAsnGlyHisSerThr 384

seq_name: pir2:B41007

seq_documentation_block:
bombersin receptor, neuromedin B-preferring - human
C:Species: Homo sapiens (man)
C:Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 16-Jul-1999
C:Accession: B41007
R:Corjay, M.H.; Dobrzanski, D.J.; Way, J.M.; Viallet, J.; Shapira, H.; Worland, P.; Saus
J. Biol. Chem. 266, 18771-18779, 1991
A:Title: Two distinct bombesin receptor subtypes are expressed and functional in human I
A:Reference number: A41007; MUID:92011639
A:Accession: B41007
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <COR>
A:Cross-references: GB:M73482; NID:g189241; PIDN:AAA59939.1; PID:g189242
C:Genetics:
A:Gene: GDB:NBWR
A:Cross-references: GDB:l28063; OMIM:162341
A:Map position: 6q21-6qter
C:Superfamily: endothelin receptor B

C;Keywords: G protein-coupled receptor; transmembrane protein
alignment_scores:
Quality: 434.50 Length: 411
Ratio: 1.849 Gaps: 13
Percent Similarity: 57.178 Percent Identity: 30.414
alignment_block:
US-09-931-157-1 x B41007 ..
Align seg 1/1 to: B41007 from: 1 to: 390

607 TGGCACAGAGCTCAGCTTCTGTTTACCACTCATCAACCCACTAAATTTGG 656
|||||
25 TrpGluArgAspPheLeuProAlaSerAspGlyThrThrThrGluLeuV 41
657 TCCTACCCAGCAATGGCTCAATGCACAACTATGCCCCAGACGACTAA 706
|||||
41 allele.....ArgCysVal 45
707 ATTACTTCAGCTTCAAAATACATTAACTGTGATATCTGTACTATTTT 756
|||||
46 IleProSerLeuTyrLeuLeuilelethr..... 55
757 CATCTGGGAATGGTGGGAATGCAACTCTGCTCAGGATCATTTACCAGA 806
|||||
56ValGlyLeuLeuGlyAsnIleMetLeuValLysIlePheIleThra 71
807 ACAATATGATAGGAATGCCCAACGCGCTGATAGCCAGTCTTGCCCTT 856
|||||
71 snSerAlaMetArgSerValProAsnIlePheIleSerAsnLeuAlaAla 87
857 GGAGACCTTATCTATGCTGCTCATTCCTCCCTATCAATGTTATTAAGCT 906
|||||
88 GlyAspLeuLeuLeuLeuThrCysValProValAsp..... 100
907 GCTGCTGGCGCTGGCTTGTGATCACAATGACTTTGGCGTATTTCTTT 956
|||||
101AlaSerArgTyrPhePheAspGluTrpMetPheGlyLysValGlyC 116
957 GCAAGCTGTTCCTCTTTTCAGAAAGTCTCGGTGGGATCACCGTCCCTC 1006
|||||
116 ysLysLeuileProValileGlnLeuThrSerValGlyValSerValPhe 132
1007 ACCCTTCGCTCTTAGTGTTCACAGGTACAGACAGTTCCTCCTCGGAG 1056
|||||
133 ThrLeuThrAlaLeuSerAlaAspArgTyrArgAlaileValAsnProm 149
1057 TCGTGTTCAGGAATGGGATTCCTTTGTAACCTGCAATGAAATGTCT 1106
|||||
149 tAspMetGlnThrSerGlyAlaLeuLeuArgThrCysValLysAlaMetG 166
1107 CCATCTGGATCTGCTCTTTATCTGGCCATTCCTGGAAGCGATT..... 1150
|||||
166 lyIleTrpValValSerValLeuLeuAlaValProGluAlaValPheSer 182
1151GGCTTC....GTCATGCT 1164
183 GluValAlaAlaArgIleSerSerLeuAspAsnSerSerPheThrAlaCysI 199
1165 ACCCTTTGAATATAGGGTGAACAGCATATAAACCTGTATGCTCAATGCCA 1214
|||||
199 eProTyrProGlnThrAspGluLeuHis..... 208
1215 CATCAAAATTCATGGAGTTCTACCAAGATGTAAGGACTGGTGGCTTTC 1264
|||||
209ProLysileHisSerValLeuilePhe 217
1265 GGGTCTATTCTGTATGCCCTTGGTGTGCTGCTGCTGCTTCTACACCT 1314
|||||
218 LeuValTyrPheLeuileProLeuAlaileleSerileTyrTyrHi 234

```
1315 CATGACTGTGAGATGTTGAACAGAGGAATGGCAGCTTGAGAAATTGCC 1364
      ::::: ::::: ::::: :::::
234 sileAlaLysThrLeuileLysSerAlaHis...AsnLeuProGlyGlu 250
      ::::: ::::: ::::: :::::
1365 TCAGTGAACATCTTAAGCAG .....CGTCGAGAAGTGGCAAAA 1402
      ::::: ::::: ::::: :::::
250 yrAsnGluHisThrLysGlnMetGluThrArgLysArgLeuAlaLys 266
      ::::: ::::: ::::: :::::
1403 ACAGTTTTCGCTGGTGAATTTTTCCTCTTTCGCTGGTTCCTCTTCA 1452
      ::::: ::::: ::::: :::::
267 ileValLeuValPheValGlyCysPheilePheCysTrpPheProAsnHi 283
      ::::: ::::: ::::: :::::
1453 CTTAAGCCGTATATTGAAGAAACTGTGTATACGAATGGACAGAACC 1502
      ::::: ::::: ::::: :::::
283 sileLeuTyrMetTyrArgSerPheAsnTyrAsnGluileAspProSerL 300
      ::::: ::::: ::::: :::::
1503 GATGTGAATTACTTAGTTTCTTACTGCTCATGGATTACATCGGTATTAA 1552
      ::::: ::::: ::::: :::::
300 euGlyHisMetileValThrLeu.....ValAlaArgVal 311
      ::::: ::::: ::::: :::::
1553 TTGGCAACCATGAATTCATGTATAAACCCCATAGCTCTGTATTTGTGAG 1602
      ::::: ::::: ::::: :::::
312 LeuSerPheGlyAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSe 328
      ::::: ::::: ::::: :::::
1603 CAAGAAATTTAAATTTGTTCCAGTCATGCCTCTGCTGC.....TGCT 1646
      ::::: ::::: ::::: :::::
328 rGluSerPheArgHisPheAsnSerGlnLeuCysCysGlyArgLys 345
      ::::: ::::: ::::: :::::
1647 GTTACCAGTCCAAAGT.....CTGATGACCTCGTCCCGCATGAAC 1687
      ::::: ::::: ::::: :::::
345 erTyrGlnGluArgGlyThrSerTyrLeuLeuSerSerSerAlaValArg 361
      ::::: ::::: ::::: :::::
1688 GGAACAAGCATCCAG.....TGGAAGAACCCAGCATCAAAACACCACAA 1731
      ::::: ::::: ::::: :::::
362 MetThrSerLeuLysSerAsnAlaLysAsnMetValThrAsnSerValLe 378
      ::::: ::::: ::::: :::::
1732 CACAGACCGGAGCGCCATAGACAGCATG 1762
      ::::: ::::: ::::: :::::
378 uLeuAsnGlyHisSerMetLysGlnGluMet 388
```

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2002, 17:15:22 ; Search time 221.29 Seconds
(without alignments)
4556.580 Million cell updates/sec

Title: US-09-931-157-1

Perfect score: 4105

Sequence: 1 gaattcgccgcctcttg.....taaaagttacagaacctt 4105

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4105	100.0	4105	4	US-08-121-446-1
2	395.8	9.6	1873	4	US-08-910-864-13
3	395.8	9.6	4301	4	US-08-121-446-3
4	101.8	2.5	1700	5	PCT-US92-02091-1
5	91	2.2	1726	5	PCT-US92-02091-3
6	79.6	1.9	1205	3	US-09-120-772-1
7	68.6	1.7	1352	5	PCT-US92-02091-7
8	65.8	1.6	1584	5	PCT-US92-02091-5
9	63.6	1.5	1563	1	US-08-279-590A-1
10	63.6	1.5	1563	1	US-08-910-092-1
11	51.4	1.3	3095	6	5231168-1
12	46.8	1.1	1981	4	US-08-387-707-15
13	46.2	1.1	595	1	US-08-784-289-1
14	46	1.1	2563	2	US-08-553-436A-7
15	45.8	1.1	19124	2	US-08-487-826B-13
16	44.8	1.1	1443	3	US-08-959-381A-3
17	44.8	1.1	1626	3	US-08-959-381A-4
18	44.2	1.1	1610	3	US-08-889-108-7
19	44.2	1.1	1610	5	PCT-US94-10358-7
20	44.2	1.1	2160	4	US-08-188-275A-1
21	44.2	1.1	2162	4	US-09-351-198-1
22	44.2	1.1	2162	4	US-09-113-426-1
23	44.2	1.1	4156	1	US-08-465-687A-1
24	44.2	1.1	4156	3	US-09-030-970-1
25	44.2	1.1	4156	5	PCT-US94-11843-1
26	43.6	1.1	1422	1	US-08-319-704-5
27	43.6	1.1	8700	2	US-08-392-625-16
28	43.6	1.1	8700	2	US-08-466-961A-16
29	43.6	1.1	8700	2	US-08-645-193B-18
30	43.4	1.1	1850	3	US-08-617-860B-32
31	43.4	1.1	2570	2	US-09-056-075-2
32	43.4	1.1	4098	2	US-08-605-106-4
33	43.4	1.1	7218	1	US-08-232-463-14
34	43.2	1.1	1080	2	US-08-103-170-3
35	43.2	1.1	1419	2	US-08-103-170-5
36	43	1.0	1080	3	US-08-602-809-3
37	43	1.0	1080	5	PCT-US95-16472-3
38	42.4	1.0	6124	4	US-08-213-419B-3
39	42	1.0	565	2	US-08-883-795A-36
40	41.8	1.0	5852	1	US-07-867-106-2
41	41.2	1.0	2463	1	US-08-370-542-1
42	41.2	1.0	2463	1	US-08-542-358-1
43	41.2	1.0	2463	3	US-09-018-351-1
44	41.2	1.0	6152	4	US-08-973-462-1
45	41	1.0	19124	2	US-08-487-826B-13

ALIGNMENTS

RESULT 1
US-08-121-446-1
; Sequence 1, Application US/08121446
; Patent No. 6313276
; GENERAL INFORMATION:
; APPLICANT: IMURA, HIROO
; APPLICANT: NAKAO, KAZUWA
; TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,446
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,684
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 29900-20324.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4105 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 485..1768
; FEATURE:
; NAME/KEY: mat_peptide
; LOCATION: 545
; US-08-121-446-1

Sequence 16, Appl
Sequence 18, Appl
Sequence 32, Appl
Sequence 2, Appl
Sequence 4, Appl
Sequence 14, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 3, Appl
Sequence 36, Appl
Sequence 2, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 1, Appl
Sequence 13, Appl

Query Match 100.0%; Score 4105; DB 4; Length 4105;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcgccgcccctcttgcgtccagagtgagtggaagtgctggagctttggagg 60
DB 1 GAATTGCGCGCCGCTCTTCTCGGTGCCAGAGTGAGTGAAGGTCTGGAGCTTTGGGAGG 60

QY 61 agacggggagaaagactgagggcgtgttctcccggaagtttcttcttctgagagcct 120
DB 61 AGACGGGGAGGAGACACTGGAGGCGTTCCTCCGGAGTTTCTTCGTGCGAGCCCT 120

QY 121 cgcgcgcgtacagtcacatcccgctgctgagcattgtgagagcggtggagagcctt 180
DB 121 CGCGCGCGCTACAGTCAATCCGCTGCTGACGANTGTGGAGAGCGGTGGAGAGCGTT 180

QY 181 catccatcccccgggtcgtcgccgggattgggtcccgagcgacacctcccgggagaa 240
DB 181 CATCCATCCCCAGCGGTGTCGCGGGGATTGGGTCCCGAGCACCTCCCGGGAGAA 240

QY 241 gcagtgcccgaggagtttctgaagccggggaagctgtgagccgaagcgccgcgcgc 300
DB 241 GCAGTGCCCGAGGAAGTTTCTGAAGCCGGGGAAGCTGTGCAGCCGGAAGCCGCCGCCG 300

QY 301 cggagccggggagacccggccacccctccgcccacccctcgcttctccggcttcctc 360
DB 301 CGGAGCCCGGGAGACACCGGCCACCTCCGCGCCACCCACCTCGCTTCTCCGGCTTCCTC 360

QY 361 tggcccaaggc 420
DB 361 TGCGCCAGGCGCGCGGAGCCCGGAGCTGCTGCGCACGCCGCGCTCCACGGTGA 420

QY 421 aaaaagtgaagggttaaaagcagcacaaagtgcataaagagatatctctcaaaattgcct 480
DB 421 AAAAAGTGAAGGTGTAAGAGCAGCACAAAGTGCAATAAGAGATATTTCTCAAAATTGCGCT 480

QY 481 caagatgaaaccccttgctcagggatccttttgctggcactggttgatgtgtaatt 540
DB 481 CAAGATGGAACCCCTTGCTCCFCAGGGCATCCCTTTTGGCTGGCATGCTGTTGGATGTGTAAT 540

QY 541 cagtgaatactcagagatcacagcaaaatctaagcaatcatgtggatgatttcaccac 600
DB 541 CAGTGAATACTCAGAGATACAGCACAAATCTAAGCAATCATGTGGATGATTTCAACAC 600

QY 601 ttttcgtygcacagctcagctcctctggtttaccactcatcaaccactaattggctct 660
DB 601 TTTTCGTGGCAGAGCTCAGCTTCCTGGTTACCACTCATCAACCCACTAATTTGGTCTCT 660

QY 661 acccagcaatggtcaatgcaactattgccacagcagactaaaattactcagcttt 720
DB 661 ACCCAGCAATGGCTCAATGCAACACTATGTGCCACAGCAGACTAAAATTACTTCAGCTTT 720

QY 721 caaatatacattcaactgtgatatctgtactattttcctcgtgggaatgggtgggaatgc 780
DB 721 CAATATACATTAACTGTGATATCTGTACTATTTCATCTGGGAATGGTGGGAATGC 780

QY 781 aactcgtcaggatcatattaccagaaacaaatgtatgagggaatggcccaacgcgcctgat 840
DB 781 AACTCTGCTCAGGATCAATTACCAAGAACAAATGTATGAGGAATGGCCCAACGCGCTGAT 840

QY 841 agcagcttgccttgagacccttatctatgtggtcatttatctccctataaatgtatt 900
DB 841 AGCCAGTCTTCCCTTGAGAGACCTTATCTATGTGGTCAATTGATTCCTCTATCAATGTATT 900

QY 901 taagctgctgctggcgctggccttttgatcacaaatgactttggcgtatttcttcttggaa 960
DB 901 TAAGCTGCTGCTGGCGCTGGCCCTTTTGATCACAAATGACTTTGGCGTATTTCTTTGCAA 960

QY 961 gctgttccccctttttgcagaagtcctcgtggggatcacccgcctcctcaacctctgcctct 1020
DB 961 GCTGTTCCTCTTTTGCAGAAAGTCTCTCGGTGGGGATCACCGTCCCTCAACCTCTCGCGCTCT 1020

QY 1021 tagtgttacaggtacagagcagttgctcctcgtgagtcggttccagggaaattgggattcc 1080
DB 1021 TAGTGTTGACAGGTACAGAGCAGTTGCCCTCTCGAGTCTGTTTACAGGAATTTGGGATTC 1080

QY 1081 tttgtaactgcgaattgaaattgtccatctggtatcctgtccttcttcttccctggccatcc 1140
DB 1081 TTTGGTAACCTGCCATTGAAATTGCTCCATCTGGATCTGCTCTTTATCTCTGGCCATTC 1140

QY 1141 tgaagcattggtcgttcattggttaccctttgaaatagggggtgacagcagataaaacctg 1200
DB 1141 TGAAGCGATTGGCTTCGTTCATGTTGACCTTTGAATATAGGGGTGAACAGCATAAACCTG 1200

QY 1201 tatgctcaatgccacatacaaaatttcattgagttctaccagaatgtaaaaggacgtggtc 1260
DB 1201 TATGCTCAATGCCACATCAAAATTCATGGAGTCTTACCAAGATGTAAAGGACTGGTGCT 1260

QY 1261 cttcggttctatttctgtagcccttgggtgactgcgactcttctctacacccctcatgac 1320
DB 1261 CTTCCGGTTCTATTCTGTATGCTGCTGGCTGGCATCTGCTTCTACACCTCATGAC 1320

QY 1321 ttgtgagatgttgaacagaagaaatgagcgttgagaattgcccctcagtgaaacattctaa 1380
DB 1321 TTGTGAGATGTTGAACAGAAGAAATGGCAGCTTGAGAATTGCCCTCAGTGAACATCTTAA 1380

QY 1381 gcagcgtcgagaagtggcaaaaacagtttctcgttggttgtaatttttgcctctcttgcgt 1440
DB 1381 GCAGGTCGAGAGTGGCAAAACAGTTTCTGCTTGGTTGTAATTTTTCCTCTTCTGCTG 1440

QY 1441 gtccctcttcaactaaagccgtatatgaagaaaactgtgtaaacgaataatggacaagaa 1500
DB 1441 GTTCCCTCTTCACTTAAAGCCGTATATTGAAGAAAACTGTGTATAACGAATGGACAGAA 1500

QY 1501 ccgagtgaattacttagtttcttactcgtcgtatgattacatcgttattcaacttggcaac 1560
DB 1501 CCGATGTGAATTACTTAGTTTCTTACTGCTCATGATFACATCGGTATTAACTTGGCAAC 1560

QY 1561 catgaattcatgtataaaccccatagctgttattttgtgagcaagaaatttaaaatctg 1620
DB 1561 CATGAATTCTATATAAACCCCATAGCTGTGTATTGTTGTGAGCAGAAAAATTTAAAAATTG 1620

QY 1621 tttccagtcagctcctcgtcgtgtgttaccagtcctcccaaaagtctgatacctcgtctcc 1680
DB 1621 TTTCCAGTCATGCCCTCTGCTGCTGTGTGTACAGTCCAAAGTCTGATGACCTCGGTCCC 1680

QY 1681 catgaacgaaacagcatccagtggaagaacacacatcaaaacacacacacagacgcg 1740
DB 1681 CATGAACGGAACACGATCCAGTGAAGAACCCACCGATCAAAACACCAACACACAGACCG 1740

QY 1741 gagcagccataaaggacagcatgaactgacccaccttagaagcactcctcgtggtactcccat 1800
DB 1741 GAGCAGCCATAAGGACAGCATGAACCTGACCACTTGAAGCACTCTCTCGGTACTCCCAT 1800

QY 1801 aatcctctcgagagaaaaaatcacaggcaactgtgactcccggaatactctctctgac 1860
DB 1801 AATCTCTCGGAGAAAAAATCACAAGGCACTGTGACTCGGGGAATCTCTCTCTGATC 1860

QY 1861 cttctcttcaattcactcccccacacccaagaagaatgcttcccaaaaccgcaagtaga 1920
DB 1861 CTTCTTCTTAAATTCACTCCCAACCCCAAGAGAAATGCTTTCCAAACCGCAAGGTAGA 1920

QY 1921 ctggtttctccaccacacacatctacgaatcgttacttctttaaattgatctaatattacata 1980
DB 1921 CTGGTTTCTCCACCACCAACATCTACGAATCGTACTCTTCTTAAATTGATCTAATTTACATA 1980

QY 1981 ttctcgtgtgtgtatttcacgactaaaaatgtgtggagctgggggagaatgaagactgtt 2040
DB 1981 TTCTCGCTGTTGTATTTCAGCATAAAAAATGGTGGAGCTGGGGGAGAAATGAAGACTGTT 2040

QY 2041 aaatgaacacacagaagatatattactacttttgcataaaatagagcttctcaagtagag 2100
DB 2041 AAATGAACACCAAGAGATATTTACTACTTTTGCATGAAATAGAGCTTTCAGTACATGG 2100


```

:
:
: TELECOMMUNICATION INFORMATION:
:
: TELEPHONE: (415) 813-5600
:
: TELEFAX: (415) 494-0792
:
: TELEX: 706141
:
: INFORMATION FOR SEQ ID NO: 3:
:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 4301 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
:
: FEATURE:
:
: NAME/KEY: CDS
: LOCATION: 238..1566
: US-08-121-446-3

```

Query Match	9.5%	Score 395.8	DB 4	Length 4301	
Best Local Similarity	65.7%	Pred. No. 1e-92			
Matches 634	Conservative	0	Mismatches 307	Indels 24	Gaps 3
Qy	689	tgccacagcagactaaattacttcagcttccaaatacatataactgtgatacttgt 748			
Db	505	TGCCAAGACCCATCGAGATCAAGGAGACTTCAAATACATCACACGGTGTGTCTCTGC 564			
Qy	749	actattttcatctcgtgggaatgggggaatgcaactctgtcaggaatcatttaccagaac 808			
Db	565	CTTGTGTTCTGCTGGGGATCATCGGGAACCTCCACACTTCTGAGAAATATCTACAAGAAC 624			
Qy	809	aaatgatgaggaatgcccacacgcgcgtgataccagcttgccttggagaccattac 868			
Db	625	AAGTGCATGGCAACGGTCCCAATATCTTGATGCCAGCTTGGCTCTGGGAGACCTGCTG 684			
Qy	869	tatgggtcattgatccctataatgtatttaagctgtggctggcgtgcgcctttt 928			
Db	685	CACATCTGCTATTGATCCCTATCAATGCTACAACTGCTGGCAGAGGACTGGCC--- 740			
Qy	929	gatacaaatgactttggcgtatttctttgcaagctgttcccccttttgcagaagtcctcg 988			
Db	741	-----ATTGTGAGCTGAGATGTAAGCTGGTGCCTTTTCATACAGAAAGCCTCC 789			
Qy	989	gtggggatcacccgtccccaacctctgcgctcttagtgtgacagggtacagagagtgccc 1048			
Db	790	GTGGGAATCACTGTGCTGAGTCTATGTGCTCTGASTATTGACATATCGAGCTGTGTCT 849			
Qy	1049	tctggagtcggtttcagggaattgggattcctttggaactgccaattgaaattgtctcc 1108			
Db	850	CTTTGGAGTAGAATTAAGGAATTTGGGGTTCCAAAATGGACAGCAGTAGAATAATGTTTTG 909			
Qy	1109	atctggaactgtcctttatccttgccattcctgaagcattgcttcgctcatggtaccc 1168			
Db	910	ATTTTGGTGGTCTCTGTGGTTCCTGGCTGCTCCCTGAAGCCATAGGTTTGTATTAATTACG 969			
Qy	1169	tttgaatataggggtgaacagcataaaacctgtatgctcaatgccc-----acatcaaaa 1222			
Db	970	ATGGACTACAAGGAAGTTATTCTCGAATCTGCTTGTCTTCTCCCGTTCAAGACAGACGT 1029			
Qy	1233	ttaatgagttctaccagaatgataaaggactgtggtcctcttcgggctctattctgtatg 1282			
Db	1030	TTTATGCGATTTTACAAGACAGCAAAAGATTGGTGGCTTTCAGCTTCTATTTTCTGCTTG 1089			
Qy	1283	cccttggtgacactgcgactcttcacacccctcatgacttgtgagatgtgaacagaagg 1342			
Db	1090	CCATTGGCCATCACTGCATTTTATTATACATAATGACCTGTGAAATGTTG---AGAAAG 1146			
Qy	1343	aatggcagcttggaattgcacctcagtgaaacatttaagcagcgctcgagaagtggcaaaa 1402			
Db	1147	AAAAAGTGGCATGCAGATTGCTTTAAATGATCACCTAAAGCAGACGCGGAAGTGGCCAAA 1206			
Qy	1403	acagttttctgcttgggttgtaatttttgccttttgcgtgttcccccttctcacttaagccgt 1462			
Db	1207	ACCGTCTTTTGGCTGGGTCCTTGTCTTTGCCCTCTGCTGGCTTCCCTTTACCTCTCAGAG 1266			
Qy	1463	atattgaagaaactgtgataaacgaaatggacagaaccgatgtggaattacttagtttc 1522			

Db	1267	ATTCTGAAGCTCACCTTTTATATCAAGATGATCCCATGATGTAACATTTTGAGCTTT	1326
Qy	1523	ttaactgctcatggattacatcgggttaacttggcaaccatgaattcatgtataaacccc	1582
Db	1327	CTGTTGGTATTGGACTATATTGGTATCAACATGGTCTCAGTCATTTCTGCATTAAACCA	1386
Qy	1583	atagctctctgattttgtgagcaagaatttaaaaattgtttccagtcacgtcctctgctgc	1642
Db	1387	ATTGCTCTGTATTGGTGGAGCAAAAGATTCAAAAACCTGCTTTAAGTCATGCTTATGCTGC	1446
Qy	1643	tgctg	1647
Db	1447	TGGTG	1451

```

RESULT      4
PC/US92-02091-1
: Sequence 1, Application PC/TUS9202091
: GENERAL INFORMATION:
: APPLICANT: Battey Jr., James F.
: APPLICANT: Corlway, Martha H.
: APPLICANT: Feldman, Richard I.
: APPLICANT: Hawkins, Richard N.
: TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Edwin P. Ching
: STREET: 1501 Harbor Bay Parkway
: CITY: Alameda
: STATE: CA
: COUNTRY: USA
: ZIP: 94501
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PC/US92/02091
: FILING DATE: 19920313
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/426,150
: FILING DATE: 24-OCT-1989
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/533,659
: FILING DATE: 05-JUN-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Ching, Edwin P.
: REGISTRATION NUMBER: 34090
: REFERENCE/DOCKET NUMBER: A-0092C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-266-7476
: TELEFAX: 415-266-7400
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1700 base pairs
: TYPE: NUCLEIC ACID
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: HYPOTHETICAL: NO
: ORIGINAL SOURCE:
: ORGANISM: Mus musculus
: CELL LINE: Swiss 3T3
: IMMEDIATE SOURCE:
: LIBRARY: Lambda GT10
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 378..1532
: PC/US92-02091-1

```


Query Match 2.5%; Score 101.8; DB 5; Length 1700;
Best Local Similarity 49.3%; Pred. No. 1.1e-16;
Matches 308; Conservative 0; Mismatches 302; Indels 15; Gaps 1;

Qy 714 cagcttcaataacattcaacactgagatcttctgtactatttctcgtgggaatggtg 773
Db 490 CGGGCTTCATCTCATCCCTGAGTTTATGGCTTATCATCTGTATAGGTCTTATTG 549
Qy 774 ggaatgcaactctgcaggatcattaccagaaacaaatgtatgaggaatgccccaaag 833
Db 550 GCAACATCAGCTCATCAAGATCTTCTGACCGTCAAGTCCATGCCAAGCGCCAAACC 609
Qy 834 cgcgtatgagcagcttgccttggagacattatctatgtgtcattgtatctccctatca 893
Db 610 TGTTCATCTCTAGCTGCTTGGGAGAGACCTGCTGCTGTGTGACATCGCGCCCTGTGG 669
Qy 894 atgtatttaagctgctgctggcctgtgcttggatcaaatgactttgctgtatttc 953
Db 670 ATGCCAGCAAGTACCTGGCTGACAGGTGGCTA-----TTTGGCAGAATTG 714
Qy 954 ttgcaagctgttccctttttgcagaaagctcgtggtgggagatcccgctcccaacctct 1013
Db 715 GCTGCAAACTGATCCCTTTATACAACTTACTTTCAGTGGGGTGTCTGTCTTACACTTA 774
Qy 1014 gcgctcttagtgtgacaggtacagagcagttgcctcctcgtgagctgtgttcagggaaattg 1073
Db 775 CGGCACTGTCAGCTCAGAGGTACAAAGCCATTGTAGGGCCATGGATATCCAGGCATCCC 834
Qy 1074 ggaattccttggtaactgcaattgaaattgtctccatctggtatcctgtcttcttcctgg 1133
Db 835 ATGCCCTGATGAAGATCTGTCTCAAGCTGCTTGTGATGGAATGTCTCTATGTTGG 894
Qy 1134 ccattcctgaagcgtattgcttgcctgctggtacccttgaataggggtgaacagcata 1193
Db 895 CAATCCAGAGCTGTGTTTCTGACCTCCACCCTTCCATGTGAAGATACCAACCAAA 954
Qy 1194 aaacgtgtatgctcaatgccacatcaaaattcaggttctaccagaatgtaaaggact 1253
Db 955 CTTTCATTAGTTGTGCCCTTACCACACTCCCAATGAGTACACCCCTAAATCCATTCCA 1014
Qy 1254 ggtgctcttcgggtcttattctgtatgcttggctgctgctgctgctgctgctgctgctg 1313
Db 1015 TGGCTTCCTTTCTGTTTTCTACGTTATCCCACTGCGCATCATCTCTGTCTACTACT 1074
Qy 1314 tcatgactgtgagatgttgaacag 1338
Db 1075 TCATTGCCCGAAATCTGATTCAGAG 1099

RESULT 5
PCT-US92-02091-3
; Sequence 3, Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Battley Jr., James F.
; APPLICANT: Corjay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US92/02091
FILING DATE: 19920313
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/426,150
FILING DATE: 24-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/533,659
FILING DATE: 05-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34090
REFERENCE/DOCKET NUMBER: A-0092C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-266-7476
TELEFAX: 415-266-7400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1726 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
CELL TYPE: Small cell lung carcinoma
CELL LINE: NCI-H345
IMMEDIATE SOURCE:
LIBRARY: Lambda GT10
FEATURE:
NAME/KEY: CDS
LOCATION: 399..1553
PCT-US92-02091-3

Query Match 2.2%; Score 91; DB 5; Length 1726;
Best Local Similarity 48.3%; Pred. No. 7.1e-14;
Matches 299; Conservative 0; Mismatches 305; Indels 15; Gaps 1;

Qy 720 tcaataacataaactgtgatacttctgtactattttctcgtgggaatggtgggaatg 779
Db 514 TCTCTATGTCATCCCTGCGAGTTTATGGGTTTATCATCTGTAGGCTCATTTGGCAACA 573
Qy 780 caactctgtcgtgagatcattaccagaaacaaatgtatgaggaatgcccccaacgcgtga 839
Db 574 TCACITTTGATCAAGATCTTCTGTACAGTCAAGTCCATGCGAAGCGTCCAAACCTGTTC 633
Qy 840 tagccagcttgccttggagacatttatctgtgtgcttgccttgccttgccttgccttgc 899
Db 634 TTTCCAGTCTGGCTTTGGGAGACCTGCTCTCTAATAAGCTGTGTCCAGTGGATGCCA 693
Qy 900 ttaagctgtgctggcgctggccttggatcaaatgactttgctgtatttcttcttga 959
Db 694 GCAGGTACCTGCTGACAGATGGCTATTGGCAGGATGGC-----TGCA 738
Qy 960 agctgttcccttttgcagaaagtcctcgtggtgggagatcacctcctcaacctctgcgctc 1019
Db 739 ACTGATCCCTTTATACAGCTTACCTCTGTGGGGTGTCTGTCTTCACACTCAGCGCGC 798
Qy 1020 ttagtgtgacaggttacagagcagttgcctcctcctggtgagtcgtgttcaggggaattgggattc 1079
Db 799 TCTGGCAGACAGATACAAAGCCATTGTCCGGCAATGGATATCCAGGCTCCCATGCC 858
Qy 1080 ctttggtaactgccattgaaattgtctccatctggtatccttgccttgccttgccttgccttgc 1139
Db 859 TGATGAAGATCTGCTCAAAAGCGCTTTTATCTGGATCATCTCCATGCTCTGGCCATTC 918
Qy 1140 ctgaagcgaattgcttgcctggtacccttgaataggggtgaacagcataaaacct 1199
Db 919 CAGAGCCCGTGTGTTTCTGTACCTCCATCCCTTCCATGAGGAAACACCAACCACTTCA 978
Qy 1200 gtatgctcaatgccacatcaaaattcattgagttctaccagaatgtaaaggactggtg 1259

Db 979 TTAGCTGTGCCCCATACCACACTCTAATGAGCTTCACCCCAAAATCCATCTATGGCTT 1038
 Qy 1260 tcttcgggtctattctgtatgcccttggtgctgactgcgatctctacacctcatga 1319
 Db 1039 CCTTCTGGTCTCTACGGTCATCCACATCGCATCTCTGTTTACTACTACTTCATG 1098
 Qy 1320 cttgtgagatggtgaacag 1338
 Db 1099 CTAAAAATCTGATCCAG 1117

```

RESULT      6
US-09-120-772-1
: Sequence 1, Application US/09120772
: Patent No. 6143521
: GENERAL INFORMATION:
: APPLICANT: LANE, PAMELA
: APPLICANT: TSUI, PING
: APPLICANT: ELSHOUBAGY, NABIL
: TITLE OF INVENTION: HUMAN BOMBESIN RECEPTOR SUBTYPE
: TITLE OF INVENTION: 3
: NUMBER OF SEQUENCES: 2
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Ratner & Prestia
: STREET: P.O. Box 980
: CITY: Valley Forge
: STATE: PA
: COUNTRY: USA
: ZIP: 19482
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/120,772
: FILING DATE: 22-JUL-1998
: CLASSIFICATION:
: PRIOR APPLICATION NUMBER:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Prestia, Paul F
: REGISTRATION NUMBER: 23,031
: REFERENCE/DOCKET NUMBER: GP-70505
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 610-407-0700
: TELEFAX: 610-407-0700
: TELEX: 946169
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1205 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cdna
US-09-120-772-1

```

Db	300	GTGTCGCCAGTGGATGCACTCACTCACTTGCAGAAGGATGGCTGTTCTC-----	347
Qy	941	tttggcgtattctctttgcaagctgttcccccttttttggaaagctcctcgttggggatcaacc	1000
Db	348	--GGAAGAAATTGGTTTAAAGTGCCTCTTTTCACTCCGGCTCACTTCTGTGGTGTCTCA	404
Qy	1001	gtcctcaacctctgcgctcttgaattgtgacaggttacagagcagttgctcctcctggagtcgt	1060
Db	405	GTGTTTCACATATTAGCAATTTCTCAGCGCTGACAGATACAAAGGCAGTTGTGAAGCCACTTTGAG	464
Qy	1061	gtcaggggaattgggattccttctgttaactgccattgaaattgtctccatctgagatcctg	1120
Db	465	CGACAGCCCTCCAAATGCCATCTCTGAAGACTTGTGTAAAGCTGGCTGGGTCTGGATCGTG	524
Qy	1121	tcctttatcctggccaatcctcgaagcattggctctgcctcgttaccctttgaaatatagg	1180
Db	525	TCTATGATATTTCCTACCTCAGGCTATATTTTCAAAATGTATACACTTTTCGAGATCCC	584
Qy	1181	ggtgaaacagcataaaacctgtatgctcaatgccacatcaaaatcattcaggtcttaccaaa	1240
Db	585	AATAAAAAATATGACATTTGAATCATGTACCTCTTATCTGTCTCTAAGAAGCTCTTCGAA	644
Qy	1241	gattgaaaggactggcgctcttcgggtctctattctctgtatgcctctgggtgcactgcg	1300
Db	645	GAATACATTCCTGCTGTGCTCTCTTAGTGTTCTACATTTATCCACTCTCTATTATCTCT	704
Qy	1301	atctctcacccctcatgacttgtagattgtgaacagaagaatggcagcttgagaatt	1360
Db	705	GTCTACTATTCTCTGATTGCTTAGGACCCCTTTACAAAAGCACCCCTGAACATACCTTACTGAG	764
Qy	1361	gcctcagtgacaatctctaagcag-----cgtcgagaggggcaaaaacagtttctc	1411
Db	765	GAAACAGCCATGCCCGCTAAGCAGATTGAATCCCGAAAGAGAAATGGCAGAACGGTATTG	824
Qy	1412	tgccttggttgtaattttgctctcttgcgtgggttccccctctcacttaagcgtatataggag	1471
Db	825	GTGTTGGTGGCTCTGTTTGGCCCTCTGCTGGTTGGCCAAATCACCT--CCTGTACCTCTACC	882
Qy	1472	aaactgtgtatacagaataatgacagaacagatggaattacttagttttcttactgctc	1531
Db	883	ATTCAATTCACTTCTCAAAAGCTATGTAGAGCCCTCTGCCATGCATTTTATTTTACCAATT	9411
Qy	1532	atggattacatcggttataactctggccaacatgaattcatgtataaaacccccatagctctg	1591
Db	942	-----TTCTCTCGGGTTTGGCTTTCACCAATTTCTTGGGTAAACCCCTTTGCTCTC	992
Qy	1592	tattttgtgagcaagaataattaaaaattgtttccagtcattgcctctctgctgcgc	1645
Db	993	TACTGGCTGAGCAAAAAGCTTTCAGAAGCATTTTAAAGCTCAGTGTCTTCTGTTCG	1046

```

RESULT 7
PCT-US92-02091-7
: Sequence 7, Application PC/TUS9202091
:
: GENERAL INFORMATION:
:
: APPLICANT: Battey Jr., James F..
:
: APPLICANT: Corjay, Martha H.
:
: APPLICANT: Feldman, Richard I.
:
: APPLICANT: Watkins, Richard N.
:
: TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
:
: NUMBER OF SEQUENCES: 8
:
: CORRESPONDENCE ADDRESS:
:
: ADDRESSEE: Edwin P. Ching
:
: STREET: 1501 Harbor Bay Parkway
:
: CITY: Alameda
:
: STATE: CA
:
: COUNTRY: USA
:
: ZIP: 94501
:
: COMPUTER READABLE FORM:
:
: MEDIUM TYPE: Floppy disk
:
: COMPUTER: IBM PC compatible
:
: OPERATING SYSTEM: PC-DOS/MS-DOS

```



```
Db 864 AAGAAATATTGAAATCTAATTAATTTGAAATAAACCATCGAAATATATTG 805
Qy 2707 cactgtatatagaagtcctaaacacacacacctaaagagaaaaagatcgaatttttcagatgatt 2766
Db 804 AAATATATAAATGAATTTATAAAATACCGAGTATTATATTATTAGGCTATTACCTTCTATT 745
Qy 2767 cggaaatttttcattcaggatttttgtaataagtgacatatatatgtatatatcacctc 2826
Db 744 AATCAATAGTTATTAAGACCTATATAAGTATATAACGACATAAATATATCGATAAAATC 685
Qy 2827 ctattctcttaatttttgtaaaagtta 2855
Db 684 GAAATTTATTTTAAATATTTTCACATGTGA 656
```

Search completed: October 1, 2002, 01:00:07
Job time: 27885 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run On: September 30, 2002, 17:06:01 ; Search time 1153.8 Seconds
(without alignments)
6108.454 Million cell updates/sec

Title: US-09-931-157-1

Perfect score: 4105

Sequence: 1 gaattcgccgcctcttg.....taaaagtgtacagaacacct 4105

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 segs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_032802.*

- 1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4103.4	100.0	4105	14	AAQ34583
2	4103	100.0	4105	21	AAA38341
3	1819.6	44.3	1868	21	AAF20903
4	1819.6	44.3	1868	21	AAF20915
5	1819.6	44.3	1868	21	AAA34781
6	1819.6	44.3	1868	21	AAA34793
7	1819.6	44.3	2008	21	AAF20904
8	1819.6	44.3	2008	21	AAA34782
9	1819.6	44.3	5036	21	AAF21447

10	1819.6	44.3	117609	21	AAF21435
11	1296.8	31.6	1310	21	AAF20902
12	1296.8	31.6	1310	21	AAF20914
13	1296.8	31.6	1310	21	AAA34780
14	1296.8	31.6	1310	21	AAA34792
15	1296.8	31.6	146981	21	AAF21442
16	1280.8	31.2	1284	23	ABI97987
17	1248.8	30.4	1284	15	AAQ63209
18	403.8	9.8	1965	13	AAQ25892
19	399	9.7	4301	14	AAQ34584
20	396.6	9.7	1406	15	AAQ53922
21	395.8	9.6	1632	24	AAQ24966
22	395.8	9.6	1719	21	AAF21285
23	395.8	9.6	1720	21	AAA35163
24	395.8	9.6	1872	21	AAF21283
25	395.8	9.6	1872	21	AAA35161
26	395.8	9.6	1873	19	AAV17875
27	395.8	9.6	4286	21	AAF21284
28	395.8	9.6	4286	21	AAA35162
29	395.8	9.6	13611	21	AAF21288
30	395.8	9.6	13612	21	AAA35166
31	392.6	9.6	1329	23	ABI97988
32	383.2	9.3	1958	24	ABI99321
33	381.2	9.3	1578	21	AAF21287
34	381.2	9.3	1578	21	AAA35165
35	343	8.4	343	20	AAH87097
36	256.2	6.2	267	21	AAQ04557
37	255	6.2	255	19	AAH11241
38	193	4.7	194	20	AAH87011
39	137.8	3.4	114955	20	AAH53491
40	128.4	3.1	784	20	AAH54750
41	128.4	3.1	784	21	AAF20319
42	128.4	3.1	784	21	AAA34197
43	128.4	3.1	7800	21	AAF20842
44	128.4	3.1	7800	21	AAA34720
45	128.4	3.1	7803	20	AAH55272

ALIGNMENTS

RESULT 1

AAQ34583

ID AAQ34583 standard; DNA; 4105 BP.

XX AAQ34583;

XX AAQ34583;

DT 11-MAY-1993 (first entry)

XX Eta receptor gene.

DE Human; ETA; ETB; endothelin; receptor; transmembrane domain; N tail;

KW extracellular; cytoplasmic; C tail; post translational; bovine;

KW modification; ET-1 receptor; antagonist; circulatory system; ss.

OS Homo sapiens.

XX Key

XX CDS

FT Location/Qualifiers

FT 485..1768

FT /tag= a

FT sig_peptide

FT 485..544

FT /tag= b

FT mat_peptide

FT 545..1765

FT /tag= c

FT polyA_signal

FT 4084..4089

FT /tag= d

FT misc_feature

FT 1972..1976

FT /tag= e

FT /function= Related with mRNA instability

FT 2059..2063

FT /tag= f

FT /function= Related with mRNA instability

FT 2309..2313

Human receptor-rel
Human endothelin r
Human ELAM-1 polyn
Human adenosine re
Human adenosine re
Human ELAM-1 polyn
Non-endogenous hum
Human endothelin r
Sequence encodes e
ETb receptor gene.
Bovine ET receptor
Human G-protein co
Human low adenosin
Human adenosine re
Human low adenosin
Human adenosine re
Homo sapiens endot
Human low adenosin
Human adenosine re
Human low adenosin
Human adenosine re
Non-endogenous hum
Mouse ischaemic co
Human low adenosin
Human adenosine re
Human single nucle
Human secreted pro
Human biallelic po
Human single nucle
Human adenosine Al
Endothelin recepto
Endothelin ETA rec
Human adenosine re
Human multiple tar
Human adenosine re
Human receptor-rel

```
FT /*tag= g
FT /function= Related with mRNA instability
FT misc_feature 2386..2390
FT /*tag= h
FT /function= Related with mRNA instability
FT misc_feature 2680..2684
FT /*tag= i
FT /function= Related with mRNA instability
FT misc_feature 3252..3256
FT /*tag= j
FT /function= Related with mRNA instability
FT misc_feature 3944..3948
FT /*tag= k
FT /function= Related with mRNA instability
XX
XX EP522868-A.
XX
XX
XX 13-JAN-1993.
XX
XX 10-JUL-1992; 92EP-0306347.
XX
XX 12-JUL-1991; 91JP-0172828.
XX
XX (SHIO ) SHIONOGI SEIYAKU KK.
XX
XX Imura H, Nakanishi S, Nakao K;
XX
XX WPI: 1993-010677/02.
XX
XX P-PSDB; AAR30885.
XX
XX Human ETA and ETB endothelin receptors - for measuring endothelin
XX and screening for endothelin antagonists
XX
XX Claim 6; Fig 1; 39pp; English.
XX
XX The sequences given in AAQ34583-84 encode the human ETA and ETB
XX endothelin receptors respectively. ETA is a 427 amino acid protein
XX with a molecular weight of 48,726. ETB comprises 442 amino acids and
XX has a molecular weight of 49,629. ETA has a higher affinity for
XX endothelin (ET)-1 and ET-2, whereas ETB has no selectivity for ET-1,
XX ET-2 or ET-3. The receptors each contain seven transmembrane domains
XX and have an extracellular N tail and a cytoplasmic C tail. There are
XX several potential sites for post translational modification, these
XX sites are identical to those of bovine ET-1 receptor. ETA cDNA is
XX 91.2% homologous to bovine ET-1 receptor cDNA and ETB cDNA is 61.1%
XX homologous to that of bovine ETA-receptor. The receptor proteins are
XX useful as reagents for measuring the amount of ET or screening for
XX antagonists of the ET receptor when studying the circulatory system.
XX
XX Sequence 4105 BP; 1138 A; 860 C; 844 G; 1263 T; 0 other;
```

```
Query Match 100.0%; Score 4103.4; DB 14; Length 4105;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 4104; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gaattcgccgcgcctcttgcggtccacagtgagtgagtgaaagctcggagcttggagg 60
DB 1 gaattcgccgcgcctcttgcggtccacagtgagtgagtgaaagctcggagcttggagg 60
QY 61 agacgggggagacagactggagcggtgttccctccgagtttttttttcgtgcgagccct 120
DB 61 agacgggggagacagactggagcggtgttccctccgagtttttttttcgtgcgagccct 120
QY 121 cgcgcgcgcgtacagtcacccgctggctggctgacgattgtggagagcggtggagagcctt 180
DB 121 cgcgcgcgcgtacagtcacccgctggctggctgacccattgtggagagcggtggagagcctt 180
QY 181 catccatcccccgcgtcgcggggttggtgggtcccgagcacacctcccccggagaa 240
DB 181 catccatcccccgcgtcgcggggttggtgggtcccgagcacacctcccccggagaa 240
QY 241 gcagtgcccgaggaaagtttctgaagccgggggaagctgtgcagccgagccgcgcgcgc 300
```

```
DB 241 gcagtgcccgaggaaagtttctgaagccggggaagctgtgcagccgaaagccgcgcgcgc 300
QY 301 cggagcccgggacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
DB 301 cggagcccgggacacgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 360
QY 361 tggcccgaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
DB 361 tggcccgaggcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
QY 421 aaaaagtgaagtgtaaaagcagcacaaagtgaataagaagatatttctcaaaattgcct 480
DB 421 aaaaagtgaagtgtaaaagcagcacaaagtgaataagaagatatttctcaaaattgcct 480
QY 481 caagatggaaaccccttgcctcagggcattcctttgctggcactggttgatgtgtaat 540
DB 481 caagatggaaaccccttgcctcagggcattcctttgctggcactggttgatgtgtaat 540
QY 541 cagtgataatcctgagagatacacagcaacaatctaagaacaatcatgtgatgatttcaccac 600
DB 541 cagtgataatcctgagagatacacagcaacaatctaagaacaatcatgtgatgatttcaccac 600
QY 601 tttctgtgcacagagctcagctcctctgtgttacactcactcaaccactaatttggctct 660
DB 601 tttctgtgcacagagctcagctcctctgtgttacactcactcaaccactaatttggctct 660
QY 661 acccgcaatggctcaatgcacaactattgccacacagactaaataacttcaagctt 720
DB 661 acccgcaatggctcaatgcacaactattgccacacagactaaataacttcaagctt 720
QY 721 caaatacattacaactgtgatatcttctgactatttctcgtgggaatggtgggaatgc 780
DB 721 caaatacattacaactgtgatatcttctgactatttctcgtgggaatggtgggaatgc 780
QY 781 aactctgctcagatcatttaccagaacaatgatgatgaggaatggcccaacgcgcgtgat 840
DB 781 aactctgctcagatcatttaccagaacaatgatgatgaggaatggcccaacgcgcgtgat 840
QY 841 agccagcttgccttggagacatttctatgtgtgctgattgctcctcattcaatgtatt 900
DB 841 agccagcttgccttggagacatttctatgtgtgctgattgctcctcattcaatgtatt 900
QY 901 taagctgctgctggcgctggccttttgcacaaactgactttgctgcttatttcttgcac 960
DB 901 taagctgctgctggcgctggccttttgcacaaactgactttgctgcttatttcttgcac 960
QY 961 gctgttcccttttttgcagaagtcctcgtgggagatcacgcgtcctcactcctgcgctct 1020
DB 961 gctgttcccttttttgcagaagtcctcgtgggagatcacgcgtcctcactcctgcgctct 1020
QY 1021 tagtggtgacaggtacagacagcttgcctcctcgtgagtcggttcaggaaattggattccc 1080
DB 1021 tagtggtgacaggtacagacagcttgcctcctcgtgagtcggttcaggaaattggattccc 1080
QY 1081 ttgtgtaactgcaattgaaattgtctccattgctgactccttcccttattcctggcattccc 1140
DB 1081 ttgtgtaactgcaattgaaattgtctccattgctgactccttcccttattcctggcattccc 1140
QY 1141 tgaagcgattggttcgtcattgacccctttgaaataaggggtgaaacagcataaaacctg 1200
DB 1141 tgaagcgattggttcgtcattgacccctttgaaataaggggtgaaacagcataaaacctg 1200
QY 1201 tatgtcctaagccacatcaaaattcatggagttctccaagatgataaaggactggtggct 1260
DB 1201 tatgtcctaagccacatcaaaattcatggagttctccaagatgataaaggactggtggct 1260
QY 1261 ctctgggttctatttctgtatgccttggctgctgactgcgactcttctacacccctcagac 1320
DB 1261 ctctgggttctatttctgtatgccttggctgctgactgcgactcttctacacccctcagac 1320
QY 1321 ttgtgagatgtgaacagaaggaaatggcagcttgagaaattgcccctcagtgaaactcttaa 1380
DB 1321 ttgtgagatgtgaacagaaggaaatggcagcttgagaaattgcccctcagtgaaactcttaa
```

Db	1321	ttgtgagatg	ttgaacagaaggaatggcagcttgagaa	tgcctcagtgaa	catcttaa	1380			
Qy	1381	gcagcgtcga	gaatggcaaaaacagtttctcgtctgg	tgtaa	tttttgcctttgctg	1440			
Db	1381	gcagcgtcga	gaatggcaaaaacagtttctcgtctgg	tgtaa	tttttgcctttgctg	1440			
Qy	1441	gttcctcttcact	ttaaagccgtatata	tgaaagaaaact	tgtataa	cgaaatggacaagaa	1500		
Db	1441	gttcctcttcact	ttaaagccgtatata	tgaaagaaaact	tgtataa	cgaaatggacaagaa	1500		
Qy	1501	ccagatgaa	tacttaagtcttctactgc	tcatggattacat	cggattaa	cttttgccac	1560		
Db	1501	ccagatgaa	tacttaagtcttctactgc	tcatggattacat	cggattaa	cttttgccac	1560		
Qy	1561	catgaattcat	gtataaaaccccatagctcgtatt	ttgtgagcaagaa	atttaaaatg	1620			
Db	1561	catgaattcat	gtataaaaccccatagctcgtatt	ttgtgagcaagaa	atttaaaatg	1620			
Qy	1621	tttccagtc	atgcctctgctgttaccag	tccaaaagtctgat	cactcgtgtccc	1680			
Db	1621	tttccagtc	atgcctctgctgttaccag	tccaaaagtctgat	cactcgtgtccc	1680			
Qy	1681	catgaacgga	caacagcatccactggagaa	ccacagatc	caaaaacacacacacacac	1740			
Db	1681	catgaacgga	caacagcatccactggagaa	ccacagatc	caaaaacacacacacacac	1740			
Qy	1741	gagcagcc	atgaagcacatgaa	ctgacaccccttaga	gcactcctcgttactcccat	1800			
Db	1741	gagcagcc	atgaagcacatgaa	ctgacaccccttaga	gcactcctcgttactcccat	1800			
Qy	1801	aatcctctcgg	gagaaaaatac	caagggcaact	tgactcgggaatctctctc	1860			
Db	1801	aatcctctcgg	gagaaaaatac	caagggcaact	tgactcgggaatctctctc	1860			
Qy	1861	cttcttctta	atttactccacaccccaagaa	aatgctttccaaa	accgcgaagttaga	1920			
Db	1861	cttcttctta	atttactccacaccccaagaa	aatgctttccaaa	accgcgaagttaga	1920			
Qy	1921	ctggtttatcc	accacacacatctacga	atcgttactctt	taattgatctaa	ttttacata	1980		
Db	1921	ctggtttatcc	accacacacatctacga	atcgttactctt	taattgatctaa	ttttacata	1980		
Qy	1981	ttctgcgtgt	gtgtattcagcactaaaa	aatggctggggagctgggg	ggagaa	tgaagactgtt	2040		
Db	1981	ttctgcgtgt	gtgtattcagcactaaaa	aatggctggggagctgggg	ggagaa	tgaagactgtt	2040		
Qy	2041	aaatgaac	acgaaggaatttactact	tttgcgatgaaaa	tagagctttc	caagtacatgg	2100		
Db	2041	aaatgaac	acgaaggaatttactact	tttgcgatgaaaa	tagagctttc	caagtacatgg	2100		
Qy	2101	ctagctttat	ggcagttcgttgaat	tccaa	tgggaactgggtccac	tatgaac	ttttag	2160	
Db	2101	ctagctttat	ggcagttcgttgaat	tccaa	tgggaactgggtccac	tatgaac	ttttag	2160	
Qy	2161	agattaac	gacagaagattttctact	ttttttaa	gtgatttttgccttc	agccaaacaca	2220		
Db	2161	agattaac	gacagaagattttctact	ttttttaa	gtgatttttgccttc	agccaaacaca	2220		
Qy	2221	atatgggtc	aggttcac	ttttttaa	tggaaatgcatt	tggcgca	gtattttttaa	2280	
Db	2221	atatgggtc	aggttcac	ttttttaa	tggaaatgcatt	tggcgca	gtattttttaa	2280	
Qy	2281	taatagc	cttaacatgat	tatttgaact	tattttacat	tagtttg	aaaaaaagacaaa	2340	
Db	2281	taatagc	cttaacatgat	tatttgaact	tattttacat	tagtttg	aaaaaaagacaaa	2340	
Qy	2341	aatag	tattcaggt	tgacaa	atttag	tattttccac	gtattattttttaa	2400	
Db	2341	aatag	tattcaggt	tgacaa	atttag	tattttccac	gtattattttttaa	2400	
Qy	2401	acacaa	atttcaag	actaca	aaatactac	agggccct	ttaagc	acagctctgat	2460
Db	2401	acacaa	atttcaag	actaca	aaatactac	agggccct	ttaagc	acagctctgat	2460

Db 301 cggagccgggaacaccgccaacctccgcgccacccaccctcgctttctcgggttcctc 360
Qy 361 tggccaggccgagccgagcccgagctgtctgcgaccccgagctccacgggtgaana 420
Db 361 tggccaggccgagccgagcccgagctgtctgcgaccccgagctccacgggtgaana 420
Qy 421 aaaaagtgaagtgtaaaagcagcacaaagtcaataagagataatttctcaaaattgctc 480
Db 421 aaaaagtgaagtgtaaaagcagcacaaagtcaataagagataatttctcaaaattgctc 480
Qy 481 caagatggaacacctttgctccaggccaccttttggctggcaactgggtggatgtgtaat 540
Db 481 caagatggaacacctttgctccaggccaccttttggctggcaactgggtggatgtgtaat 540
Qy 541 cagtgataatcctgagagatcacagacaaaactaagcaaatcatgtggatgttccacac 600
Db 541 cagtgataatcctgagagatcacagacaaaactaagcaaatcatgtggatgttccacac 600
Qy 601 tttctgtgacagagctcagcttctctgtgttaccactcatcaacccactaatttggctct 660
Db 601 tttctgtgacagagctcagcttctctgtgttaccactcatcaacccactaatttggctct 660
Qy 661 acccagcaatggctcaatgacaaactattgccacagcagactaaaattacttcagcttt 720
Db 661 acccagcaatggctcaatgacaaactattgccacagcagactaaaattacttcagcttt 720
Qy 721 caaatacataacactgtgatcttctgtactatttctcgtggaaatgggtgggaaatgc 780
Db 721 caaatacataacactgtgatcttctgtactatttctcgtggaaatgggtgggaaatgc 780
Qy 781 aactctgtcagagctattaccagaaacaaatgtatgaggaaatggccccacgcgctgat 840
Db 781 aactctgtcagagctattaccagaaacaaatgtatgaggaaatggccccacgcgctgat 840
Qy 841 agccagcttctcccttgagacccttatctatgtggatcattgatccctatacaatgtatt 900
Db 841 agccagcttctcccttgagacccttatctatgtggatcattgatccctatacaatgtatt 900
Qy 901 taagctgtcgtggcgctggccttttgatacaaatgactttggcgtatttctttgcaa 960
Db 901 taagctgtcgtggcgctggccttttgatacaaatgactttggcgtatttctttgcaa 960
Qy 961 gctgttcccttttttgagaagctcctcggtgggagataccgtcctcaacccctcggcctc 1020
Db 961 gctgttcccttttttgagaagctcctcggtgggagataccgtcctcaacccctcggcctc 1020
Qy 1021 tagtgtgacaggtacagagcagttgcctcctggagtcgtgtcagggaattgggattcc 1080
Db 1021 tagtgtgacaggtacagagcagttgcctcctggagtcgtgtcagggaattgggattcc 1080
Qy 1081 tttggtaactgccattgaaattgtctccatctggaatcctgtcttatacctggccattcc 1140
Db 1081 tttggtaactgccattgaaattgtctccatctggaatcctgtcttatacctggccattcc 1140
Qy 1141 tgaagcgattggcttcgtcatggtaccctttgaaatatagggggtgaacggcataaaacctg 1200
Db 1141 tgaagcgattggcttcgtcatggtaccctttgaaatatagggggtgaacggcataaaacctg 1200
Qy 1201 tatgctcaatgccacatacaaaatcatggagttctaccagaatgtaaaaggactgggtgct 1260
Db 1201 tatgctcaatgccacatacaaaatcatggagttctaccagaatgtaaaaggactgggtgct 1260
Qy 1261 cttcgggttctatttctgtatgccccttgggtgcaactcgatcttctacacccctcatgac 1320
Db 1261 cttcgggttctatttctgtatgccccttgggtgcaactcgatcttctacacccctcatgac 1320
Qy 1321 ttgtgagatgttgaacagaaggaatggcagcttgagaattgccctcagtgaaactcttaa 1380
Db 1321 ttgtgagatgttgaacagaannaaatggcagcttgagaattgccctcagtgaaactcttaa 1380
Qy 1381 gcagcgtcgagaagtggaacacagtttctcgtcgtgtgtatatttctcgtcttctgtg 1440
Db 1381 gcagcgtcgagaagtggaacacagtttctcgtcgtgtgtatatttctcgtcttctgtg 1440

Qy 1441 gttcccttctcaacttaagccgtatattgaaagaaaactgtataacgaaaatggacaagaa 1500
Db 1441 gttcccttctcaacttaagccgtatattgaaagaaaactgtataacgaaaatggacaagaa 1500
Qy 1501 ccgagtgaattacttagtttctactactcctcatgattacatcggtatttaacttggcaac 1560
Db 1501 ccgagtgaattacttagtttctactactcctcatgattacatcggtatttaacttggcaac 1560
Qy 1561 catgaattcatgtataaaccccatagctctgtatttttgagcaagaaatttaaaaattg 1620
Db 1561 catgaattcatgtataaaccccatagctctgtatttttgagcaagaaatttaaaaattg 1620
Qy 1621 tttccagtcagctcctctgctgtctgttaccagtcacaaagctcgaagctcgtctccc 1680
Db 1621 tttccagtcagctcctctgctgtctgttaccagtcacaaagctcgaagctcgtctccc 1680
Qy 1681 catgaacggaacaagcatccagtggaagaacacagatcaaaacacacacacacagacg 1740
Db 1681 catgaacggaacaagcatccagtggaagaacacagatcaaaacacacacacacagacg 1740
Qy 1741 gagcagcctaagcagacatgaactgacccaccttagaagcactcctcgttactcccat 1800
Db 1741 gagcagcctaagcagacatgaactgacccaccttagaagcactcctcgttactcccat 1800
Qy 1801 aatcctctcggagaaaaaaatcaagaagcaactgtgactccgggaaatcttctctctgac 1860
Db 1801 aatcctctcggagaaaaaaatcaagaagcaactgtgactccgggaaatcttctctctgac 1860
Qy 1861 cttctctcttaattcaactccacacacaaagaagaatcttccaaaaacgcaagtgaga 1920
Db 1861 cttctctcttaattcaactccacacacaaagaagaatcttccaaaaacgcaagtgaga 1920
Qy 1921 ctgggtttaccacccacacacatctcgaatcgactcttcttaattgatctaattacata 1980
Db 1921 ctgggtttaccacccacacacatctcgaatcgactcttcttaattgatctaattacata 1980
Qy 1981 tctcgcgtgtgtatctcagcactaaaaatgggtggagctgggtggagaaatggagactgtt 2040
Db 1981 tctcgcgtgtgtatctcagcactaaaaatgggtggagctgggtggagaaatggagactgtt 2040
Qy 2041 aatgaaacccagaagatatitactactcttctgcatgaaaaatagagcttctcaagtacatg 2100
Db 2041 aatgaaacccagaagatatitactactcttctgcatgaaaaatagagcttctcaagtacatg 2100
Qy 2101 ctagcttttatggcagttctcgtgaaatgttcaatgggaactggtccaccatgaaactttag 2160
Db 2101 ctagcttttatggcagttctcgtgaaatgttcaatgggaactggtccaccatgaaactttag 2160
Qy 2161 agattaaacgaagattttctactcttcttaagtgtattttgtccttcagcccaacaca 2220
Db 2161 agattaaacgaagattttctactcttcttaagtgtattttgtccttcagcccaacaca 2220
Qy 2221 atagggctcaggtcacttttatttgaatgtcatttggcgcaatttttttaactgca 2280
Db 2221 atagggctcaggtcacttttatttgaatgtcatttggcgcaatttttttaactgca 2280
Qy 2281 taatagcctaacatgattatttgaacttatttacacatagtttgaaaaaaaagacaaa 2340
Db 2281 taatagcctaacatgattatttgaacttatttacacatagtttgaaaaaaaagacaaa 2340
Qy 2341 aatagttatccaggtgagcaattagatttagttttccacgctcaactatttttttttaa 2400
Db 2341 aatagttatccaggtgagcaattagatttagttttccacgctcaactatttttttttaa 2400
Qy 2401 acacaaaattcaaaagctacacaaaatactacagggcccttaaaagcagctcgtatgacaca 2460
Db 2401 acacaaaattcaaaagctacacaaaatactacagggcccttaaaagcagctcgtatgacaca 2460
Qy 2461 tttggcagtttaataagatgttactcaagaattttttaagaactgtattttatttttaa 2520
Db 2461 tttggcagtttaataagatgttactcaagaattttttaagaactgtattttatttttaa 2520

XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
PS Disclosure; Page 676; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAR18434 to AAR21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 1868 BP; 512 A; 419 C; 389 G; 548 T; 0 other;

Query Match	44.3%	Score 1819.6;	DB 21;	Length 1868;
Best Local Similarity	99.7%;	Pred. No. 0;		
Matches 1844;	Conservative	0;	Mismatches	4;
			Indels	2;
			Gaps	2;

Qy	421	aaaaagtgaagggtgataaaagcagcaacagtgcaataaagagataatttctccaataatttgct	480
Db	10		69
		aaaaagtgaagggtgataaaagcagcaacagtgcaataaagagataatttctccaataattgct	
Qy	481	caagatggaacaccttgcctcagggcattctttggctggcagctgggtggagtggtgaat	540
Db	70		129
		cagagatggaaacctttgtcctcagggcattctttggctggcagctgggtggagtggtgaat	
Qy	541	cagtgataatcctgagagatacagcacaatactaaagcaatcatgtggatgatttcacac	600
Db	130		189
		cagtgataatcctgagagatacagcacaatactaaagcaatcatgtggatgatttcacac	
Qy	601	ttttcgtggcaagagctcagcttctcgtgtaccactcatcaaccacataatttggctct	660
Db	190	ttttcgtggcagagagctcagcttctcgtgtaccactcatcaaccacataatttggctct	249
Qy	661	accagcaatgggtcgaatgcacaactattgcccacagcagactaaaattacttcagcttt	720
Db	250		309
		accagcaatgggtcgaatgcacaactattgcccacagcagactaaaattacttcagcttt	
Qy	721	caaatacattaacactgtgatatcttctactattttctatcgtgggaatgggtgggaatgc	780
Db	310	caaatacattaacactgtgatatcttctactattttctatcgtgggaatgggtgggaatgc	369
Qy	781	aacttgcctcaggatcatattaccagaacaaaatgtagaagaatggcccccaacgcgcgtgat	840
Db	370	aacttgcctcaggatcatattaccagaacaaaatgtagaagaatggcccccaacgcgcgtgat	429
Qy	841	agccagctctggcctctggagaaccttattcatatgggtcatttgatctctccctcatcaatgatt	900

Db 610 tagtctgacaggtacagagcagttgctcctcctgagtcgtgttcaaggaaattgggattcc 669
Qy 1081 ttggttaacgcatgaaattgctccatctggtatcctcttattacctggccattcc 1140
Db 670 ttggttaacgcatgaaattgctccatctggtatcctcttattacctggccattcc 729
Qy 1141 tgaagcagattggtctgctggtaccccttgaataataggggtgaaacagcataaaacctg 1200
Db 730 tgaagcagattggtctgctggtaccccttgaataataggggtgaaacagcataaaacctg 789
Qy 1201 tatgtcaatgcacatacaaaattcatgaggttctcaccagatgtaagagcaggtggtct 1260
Db 790 tatgtcaatgcacatacaaaattcatgaggttctcaccagatgtaagagcaggtggtct 849
Qy 1261 ctctgggtcttattctgtatgctctggtggtgacactgcgatctctacacccctcatgac 1320
Db 850 ctctgggtcttattctgtatgctctggtggtgacactgcgatctctacacccctcatgac 909
Qy 1321 ttgtgagatgttgaaacagaaggaaatggcagcttgagaaatggcctcagtgaaacatcttaa 1380
Db 910 ttgtgagatgttgaaacagaaggaaatggcagcttgagaaatggcctcagtgaaacatcttaa 969
Qy 1381 gcagcgtcgagaagtgcaaaaacagtttctctgctgtgtgtaatttttgcctcttctgctg 1440
Db 970 gcagcgtcgagaagtgcaaaaacagtttctctgctgtgtgtaatttttgcctcttctgctg 1029
Qy 1441 gtctccctctcacttaagccgtatattgaaagaaactgtgtataacgaaatgggacaagaa 1500
Db 1030 gtctccctctcacttaagccgtatattgaaagaaactgtgtataacgagatggacaagaa 1089
Qy 1501 ccgatgtgaattacttagtttcttactgctcagtgatgattacatcggtattaaacttggcaac 1560
Db 1090 ccgatgtgaattacttagtttcttactgctcagtgatgattacatcggtattaaacttggcaac 1149
Qy 1561 catgaattcatgataaaaccccatagctctgtattttgtgagcaagaaattttaaataatg 1620
Db 1150 catgaattcatgataaaaccccatagctctgtattttgtgagcaagaaattttaaataatg 1209
Qy 1621 ttccagtcagctctgctgctgctgttaccagtcacaaagtctgagacctcggtctcc 1680
Db 1210 ttccagtcagctctgctgctgctgttaccagtcacaaagtctgagacctcggtctcc 1269
Qy 1681 catgaacggaacagatccagtggaagaaacccagatcaaaacacacacacacagacagc 1740
Db 1270 catgaacggaacagatccagtggaagaaacccagatcaaaacacacacacacagacagc 1329
Qy 1741 gacgagcctaagacagcatgaactgaccaccccttagagcactctcggtactcccat 1800
Db 1330 gacgagcctaagacagcatgaactgaccaccccttagagcactctcggtactcccat 1389
Qy 1801 aatctctcgagaaaaaatcaaggaactgtgactcgcgggaattcttctctgac 1860
Db 1390 aatctctcgagaaaaaatcaaggaactgtgactcgcgggaattcttctctgac 1449
Qy 1861 ctctctcttaattcaatccacacccaagaagaaatgctttccaaacacgcaa-ggtag 1919
Db 1450 ctctctcttaattcaatccacacccaagaagaaatgctttccaaacacgcaaaggtag 1509
Qy 1920 actggtttatccacccaacatcatcgaatctgactcttcttattgattgattcaatttcat 1979
Db 1510 actggtttatccacccaacatcatcgaatctgactcttcttattgattgattcaatttcat 1569
Qy 1980 attctcggtgtgtattcagcactaaaaaatggtggagctggggagaaatgaagactgt. 2039
Db 1570 attctcggtgtgtattcagcactaaaaaatggtggagctggggagaaatgaagactgt 1629
Qy 2040 taaatgaacacagaagatatattactcttttgcatagaataagatttcaagatacatg 2099
Db 1630 taaatgaacacagaagatatattactcttttgcatagaataagatttcaagatacatg 1689
Qy 2100 gctagcttttatgacgttctggtgaatttcaatggaaactggtcaccatgaacttta 2159
Db 1690 gctagcttttatgacgttctggtgaatttcaatggaaactggtcaccatgaacttta 1749

Qy 2160 gagattaacacagaagattttctactctttttaagta-tttttgtctctcagcaaca 2218
Db 1750 gagattaacacagaagattttctactctttttaagtgattttttgtctcagcaaca 1809
Qy 2219 caatatgggtcaggtcacttttatttgaaatgcatttggtgccagat 2268
Db 1810 caatatgggtcaggtcacttttatttgaaatgcatttggtgccagat 1859

RESULT 5

AAA34781
ID AAA34781 standard; DNA; 1868 BP.
XX
AC AAA34781;
DT 28-JUL-2000 (first entry)
XX
DE Human adenosine receptor related polynucleotide SEQ ID NO:2470.
XX
KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
phosphorothioate; impaired respiration; inflammation; allergy;
allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
respiratory distress syndrome; pain; cystic fibrosis; emphysema;
pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
XX
OS Homo sapiens.
XX
PN W0200009525-A2.
XX
PD 24-FEB-2000.
XX
PF 03-AUG-1999; 99WO-US17712.
XX
PR 03-AUG-1998; 98US-0095212.
XX
PA (UYEC-) UNIV EAST CAROLINA.
XX
PI Nyce JW;
XX
DR WPI: 2000-205971/18.
XX
PT New antisense oligonucleotides useful for treating e.g. pulmonary
vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers
XX
PS Disclosure: Page 606-607; 1343pp; English.

XX
XX
CC The present invention describes a new composition comprising an
antisense oligonucleotide (ON) with low adenosine (up to 15%), which
targets nucleic acids involved in bronchoconstriction, allergies, and/or
inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, cystic
asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent
CC nucleotide sequences given in the sequence listing from the present
invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
differ from the previously named sequences. SEQ ID NO:11 to 1680

CC (AAA32323 to AAA33992) are specifically claimed ONS from the present
CC invention. N.B. sequences given in the disclosure of the present
CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.
XX
XX · Sequence 1868 BP; 512 A; 419 C; 389 G; 548 T; 0 other;
SQ

SQ · Sequence 1868 BP; 512 A; 419 C; 389 G; 548 T; 0 other;

Query Match 44.3%; Score 1819.6; DB 21; Length 1868;
Best Local Similarity 99.7%; Pred. No., 0;
Matches 1844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy	421	aaaaagtgaaggtgtgtaaaaagcagcaagtcgcaataagagatatcttctcaaatgtgcct	480
Db	10	aaaaagtgaaggtgtgtaaaaagcagcaagtcgcaataagagatatcttctcaaatgtgcct	69
Qy	481	caagatggaaaccccttggctcagggcattcccttttggctggcactggttgagtgtgaat	540
Db	70	caagaaggaaaccccttggctcagggcattcccttttggctggcactggttgagtgtgaat	129
Qy	541	cagtgtataatcctgagagatacagacaaaactaaagcaatcatgtgagtattccaccac	600
Db	130	cagtgtataatcctgagagatacagacaaaactaaagcaatcatgtgagtattccaccac	189
Qy	601	ttttctggcacagagctcagcttctctggttacactcatcaaccactaatttggtcct	660
Db	190	ttttctggcacagagctcagcttctctggttacactcatcaaccactaatttggtcct	249
Qy	661	accacgaatggtcctaatagcacaactattgccacagcagactaaaaattacttcagcttt	720
Db	250	accacgaatggtcctaatagcacaactattgccacagcagactaaaaattacttcagcttt	309
Qy	721	caaatacattaaacactgtatctctgtactattttcatctggggaatggtggggaatgc	780
Db	310	caaatacattaaacactgtatctctgtactattttcatctggggaatggtggggaatgc	369
Qy	781	aactctgctcaggatcatttacagaaacaaatgtatgagggaatggcccaacgcctgat	840
Db	370	aactctgctcaggatcatttacagaaacaaatgtatgagggaatggcccaacgcctgat	429
Qy	841	agccagcttgccttggagacattatcatgtggttcattgatctccctcataatgtatt	900
Db	430	agccagcttgccttggagacattatcatgtggttcattgatctccctcataatgtatt	489
Qy	901	taagctgctcggtggcgctggcccttttgatcacaaatgaactttggcgattctcttgcaa	960
Db	490	taagctgctcggtggcgctggcccttttgatcacaaatgaactttggcgattctcttgcaa	549
Qy	961	gctgttccccctttttgcagaagctctcgtgggggataccctcctcaacctgcgctct	1020
Db	550	gctgttccccctttttgcagaagctctcgtggggataccctcctcaacctgcgctct	609
Qy	1021	tagtgtacaggttacagacagtttgctcctcggagtcgtgttcaggggaattgggattcc	1080
Db	610	tagtgtacaggttacagacagtttgctcctcggagtcgtgttcaggggaattgggattcc	669
Qy	1081	tttggtaactgccattgaaattgctcccatctggatcctgctcttattcctggccattcc	1140
Db	670	tttggtaactgccattgaaattgctcccatctggatcctgctcttattcctggccattcc	729
Qy	1141	tgaagcgattggcttgctcatggtgaccctttgaaatataggggtgaaacagcataaaacctg	1200
Db	730	tgaagcgattggcttgctcatggtgaccctttgaaatataggggtgaaacagcataaaacctg	789
Qy	1201	tatgctcaatgccacaatacaaaattcatggagttctaccaagatgtaaaggactgggtgct	1260
Db	790	tatgctcaatgccacaatacaaaattcatggagttctaccaagatgtaaaggactgggtgct	849
Qy	1261	cttcggagttctatttctgtatgcccttgggtgcaactgcgactcttcacaccctcatgac	1320
Db	850	cttcggagttctatttctgtatgcccttgggtgcaactgcgactcttcacaccctcatgac	909
Qy	1321	tttgtgagatgttgaacagaaggaatgggcagctttgagaattgccctcagtgaaacatttaa	1380

RESULT

RESUL
AAA34793

AAA34793
ID AAA34793 standard; DNA; 1863 BP.

XX
XX

AC AAA34793;

XX

DT 28-JUL-2000 (first entry)

XX
DE Human adenosine receptor related polynucleotide SEQ ID NO:2482.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.
 XX Homo sapiens.
 XX WO200009525-A2.
 PN 24-FEB-2000.
 XX 03-AUG-1999; 99WO-US17712.
 XX 03-AUG-1998; 98US-0095212.
 PR (UYEC-) UNIV EAST CAROLINA.
 XX NYce JW;
 XX WPI: 2000-205971/18.
 DR New antisense oligonucleotides useful for treating e.g. pulmonary
 XX vasoconstriction, inflammation, allergies, asthma, hypertension,
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers.
 XX Disclosure; Page 645-646; 1343pp; English.
 XX The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA35312 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1680
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.
 XX Sequence 1868 BP; 512 A; 419 C; 389 G; 548 T; 0 other;

Query Match 44.3%; Score 1819.6; DB 21; Length 1868;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 421 aaaagtgaaggtgtaaaagcagcacaaagtgcataaagagatatttctcaaaattgctt 480
 Db 10 aaaagtgaaggtgtaaaagcagcacaaagtgcataaagagatatttctcaaaattgctt 69
 Qy 481 caaatggaacacctttgctcctcagggaatccttttggctggcactggttgatgtaatt 540
 Db 70 caaatggaacacctttgctcctcagggaatccttttggctggcactggttgatgtaatt 129

Qy 541 cagtgaatactctgagagatacagcacaaatcttaagcaatcatgtgagatttccacac 600
 Db 130 cagtgaatactctgagagatacagcacaaatcttaagcaatcatgtgagatttccacac 189
 Qy 601 tttctggtgacagagctcagctcctctggtttaccactcatcaaccactaatttggctct 660
 Db 190 tttctggtgacagagctcagctcctctggtttaccactcatcaaccactaatttggctct 249
 Qy 661 acccagcaatggctcaatgcacaaactattgccacagcagactaaaaataactcagcttt 720
 Db 250 acccagcaatggctcaatgcacaaactattgccacagcagactaaaaataactcagcttt 309
 Qy 721 caaatacattaaactgtgatcttctgtactattttctatctgttgggaatggtgggaaatgc 780
 Db 310 caaatacattaaactgtgatcttctgtactattttctatctgttgggaatggtgggaaatgc 369
 Qy 781 aactctgtcaggaatcattaccagaacaaatgtatgaggaaatgccccaaacgcgcgat 840
 Db 370 aactctgtcaggaatcattaccagaacaaatgtatgaggaaatgccccaaacgcgcgat 429
 Qy 841 agccagttctgccccttgagacaccttatctatgtgtcattgatctccctatcaatgtatt 900
 Db 430 agccagttctgccccttgagacaccttatctatgtgtcattgatctccctatcaatgtatt 489
 Qy 901 taagctgctggctggcgctggcctttttgtgtacaaatgaactttggcgtattctttgcaa 960
 Db 490 taagctgctggctggcgctggcctttttgtgtacaaatgaactttggcgtattctttgcaa 549
 Qy 961 gctattcccccttttgcagaagctcctcgttgggaatcaccgtcctcaaccctcgcctct 1020
 Db 550 gctattcccccttttgcagaagctcctcgttgggaatcaccgtcctcaaccctcgcctct 609
 Qy 1021 tagtgtgacaggtacagagcagttgctcctcgtgagtcgtgttcagggaatgggattcc 1080
 Db 610 tagtgtgacaggtacagagcagttgctcctcgtgagtcgtgttcagggaatgggattcc 669
 Qy 1081 tttgtaactgccaattgaattgtctccatctgtgatctgctcttaccctggccattcc 1140
 Db 670 tttgtaactgccaattgaattgtctccatctgtgatctgctcttaccctggccattcc 729
 Qy 1141 tgaagcattggtctcctcgttgcacctttgaaataggggtgaaacagcaataaacctg 1200
 Db 730 tgaagcattggtctcctcgttgcacctttgaaataggggtgaaacagcaataaacctg 789
 Qy 1201 tatgtcgaatgccacatcaaaatctcatggagttcttaccagaatgtaaaggagctgtgct 1260
 Db 790 tatgtcgaatgccacatcaaaatctcatggagttcttaccagaatgtaaaggagctgtgct 849
 Qy 1261 cttcgggtctattctgtatgccttgggtggaactgcgattcttaccacctcatgac 1320
 Db 850 cttcgggtctattctgtatgccttgggtggaactgcgattcttaccacctcatgac 909
 Qy 1321 ttgtgagatgttgaacagaagaatgagcagcttgagaattgccctcagtgaaactctcaa 1380
 Db 910 ttgtgagatgttgaacagaagaatgagcagcttgagaattgccctcagtgaaactctcaa 969
 Qy 1381 gcagcgtcagaagtggcaaaaacagttttctgtggtgtgtaatttttgcctcttgcgtg 1440
 Db 970 gcagcgtcagaagtggcaaaaacagttttctgtggtgtgtaatttttgcctcttgcgtg 1029
 Qy 1441 gttccctcttcaacttaagccgtatattgaagaaaaactgtgtataacgaattggacaagaa 1500
 Db 1030 gttccctcttcaacttaagccgtatattgaagaaaaactgtgtataacgaattggacaagaa 1089
 Qy 1501 ccgagtgaattacttagtttcttactctcatggtacatcgtgtattacttggcgaac 1560
 Db 1090 ccgagtgaattacttagtttcttactctcatggtacatcgtgtattacttggcgaac 1149
 Qy 1561 catgaattcatgataaaaccccatagctctgtattttgtgagcaagaatttaaaaattg 1620
 Db 1150 catgaattcatgataaaaccccatagctctgtattttgtgagcaagaatttaaaaattg 1209

QY 721 caaatatcaataactgtgatatctctgtactatttctaatctgtggaatggtgggaatgc 780
|||||
Db 450 caaatatcaataactgtgatatctctgtactatttctaatctgtggaatggtgggaatgc 509
QY 781 aactctgctcagatcatattaccagaacaaaatgatgagaaatggcccaacgcgtgat 840
|||||
Db 510 aactctgctcagatcatattaccagaacaaaatgatgagaaatggcccaacgcgtgat 569
QY 841 agccagctctgcccctggagacccttatctatgtggttcattgatctccctatcaatgtatt 900
|||||
Db 570 agccagctctgcccctggagacccttatctatgtggttcattgatctccctatcaatgtatt 629
QY 901 taagctgctggtcgggcgtgcctcttttgatcaacaatgaactttggcgattttctttgcaa 960
|||||
Db 630 taagctgctggtcgggcgtgcctcttttgatcaacaatgaactttggcgattttctttgcaa 689
QY 961 gcctgtccccccttttgcagaagctccctggtggggatcaaccgtcctcaacctctgcgcctc 1020
|||||
Db 690 gcctgtccccccttttgcagaagctccctggtggggatcaaccgtcctcaacctctgcgcctc 749
QY 1021 tagtgttgacaggtacagagcagttgcctcctctggagtcgttggtcagggaattgggattcc 1080
|||||
Db 750 tagtgttgacaggtacagagcagttgcctcctctggagtcgttggtcagggaattgggattcc 809
QY 1081 ttctggtaactgcattgaaattgtctccatctggatcctctgcctttatctcctggccattcc 1140
|||||
Db 810 ttctggtaactgcattgaaattgtctccatctggatcctctgcctttatctcctggccattcc 869
QY 1141 tgaagcgattggtctgcgtcatgtaccccttgaatatagggtggaacagacataaaaacctg 1200
|||||
Db 870 tgaagcgattggtctgcgtcatgtaccccttgaatatagggtggaacagacataaaaacctg 929
QY 1201 tatgtcctaatgcccacatcaaaaattcatggagttctaccagaatgtaaagactggtggct 1260
|||||
Db 930 tatgtcctaatgcccacatcaaaaattcatggagttctaccagaatgtaaagactggtggct 989
QY 1261 ctctcgggtctattctgtatgcctctggtgtgcaactgcgatctcttaaccacctcatgac 1320
|||||
Db 990 ctctcgggtctattctgtatgcctctggtgtgcaactgcgatctcttaaccacctcatgac 1049
QY 1321 ttgtgagatgtgaacagaggaatggaacttgagaattgcccctcagtgaaacatcttaa 1380
|||||
Db 1050 ttgtgagatgtgaacagaggaatggaacttgagaattgcccctcagtgaaacatcttaa 1109
QY 1381 gcagcgtcagagagtggaacaaacagtttctctggttggtgtaatttttgcctcttgcgtg 1440
|||||
Db 1110 gcagcgtcagagagtggaacaaacagtttctctggttggtgtaatttttgcctcttgcgtg 1169
QY 1441 gtccctcttcaacttaagccgtatattgaagaaaactgtgtataacagaaatggacaaagaa 1500
|||||
Db 1170 gttccctcttcatttaagccgtatattgaagaaaactgtgtataacagagatggacaaagaa 1229
QY 1501 ccgatgtgaattacttagtttcttactgcccattggaattacatcggtattaaacttgcaac 1560
|||||
Db 1230 ccgatgtgaattacttagtttcttactgcccattggaattacatcggtattaaacttgcaac 1289
QY 1561 catgaattcatgtataaaaccccatagctctgtattttgtgagcaagaaattttaaataatg 1620
|||||
Db 1290 catgaattcatgtataaaaccccatagctctgtattttgtgagcaagaaattttaaataatg 1349
QY 1621 ttccagtcgctcctgtcgtctgttaccagtcacaaagtctgatgacctcgtctcc 1680
|||||
Db 1350 ttccagtcgctcctgtcgtctgttaccagtcacaaagtctgatgacctcgtctcc 1409
QY 1681 catgaacggaacacagcatccagtggaagaccacgcgatcaaaaacacacacagaccg 1740
|||||
Db 1410 catgaacggaacacagcatccagtggaagaccacgcgatcaaaaacacacacagaccg 1469
QY 1741 gacagccataagacagcatgaactgacacaccttagaagcactcctcgttactccccat 1800
|||||
Db 1470 gacagccataagacagcatgaactgacacaccttagaagcactcctcgttactccccat 1529
QY 1801 aatcctctcgggagaaaaaatcacaaaggcaactgtgactccgggaatctctctctgac 1860

|||||
Db 1530 aatcctctcgggagaaaaaatcacaaaggcaactgtgagtcggaatactctctctgac 1589
QY 1861 ctctctcttaattcactccacaccccaagaagaatactctccaaacccgcaa-ggtag 1919
|||||
Db 1590 ctctctcttaattcactccacaccccaagaagaatactctccaaacccgcaagggtag 1649
QY 1920 actgggtttatcccccacacacatctacgaatcgtactcttcttaattgatctaatattcat 1979
|||||
Db 1650 actgggtttatcccccacacacatctacgaatcgtactcttcttaattgatctaatattcat 1709
QY 1980 attctcgtgtgttattcagcactaaaaatggtggagctggtgggagagatgaagactgt 2039
|||||
Db 1710 attctcgtgtgttattcagcactaaaaatggtggagctggtgggagagatgaagactgt 1769
QY 2040 taaatgaacccaggaaggtatttactacttttgcatagaataagagctttcaagctacat 2099
|||||
Db 1770 taaatgaacccaggaaggtatttactacttttgcatagaataagagctttcaagctacat 1829
QY 2100 gctagcttttatggcagttctcgtgaatgttcaatgggaactggtcaccatgaaacttta 2159
|||||
Db 1830 gctagcttttatggcagttctcgtgaatgttcaatgggaactggtcaccatgaaacttta 1889
QY 2160 gagattaacgacagagattttctactcttttttaagtga-ttttttgcctcagccaaaca 2218
|||||
Db 1890 gagattaacgacagagattttctactcttttttaagtgtatttttgcctcagccaaaca 1949
QY 2219 caatatgggctcaggctcaacttttatttgaatgtcatttgggtgccagat 2268
|||||
Db 1950 caatatgggctcaggctcaacttttatttgaatgtcatttgggtgccagat 1999

RESULT 8

AAA34782

ID AAA34782 standard; DNA; 2008 BP.

XX AAA34782;

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide SEQ ID NO:2471.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
XX phosphorothioate; impaired respiration; inflammation; allergy;
XX allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
XX antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
XX lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
XX respiratory distress syndrome; pain; cystic fibrosis; emphysema;
XX pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
XX cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US17712.

XX 03-AUG-1998; 98US-0095212.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI: 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
XX vasoconstriction, inflammation, allergies, asthma, hypertension,
XX bronchitis, emphysema, respiratory distress syndrome, ischemia or
XX cancers -

XX Disclosure; Page 607; 1343pp; English.

Qy 2100 gctagcttttgcagttctcgtggaatgttcaatgggaactgtccaccatgaacttta 2159
Db 1830 gctagcttttgcagttctcgtggaatgttcaatgggaactgtccaccatgaacttta 1889
Qy 2160 gagattaacgacaagattttctacttttttaagtga-tttttgtccttcagcaaca 2218
Dy 1890 gagattaacgacaagattttctacttttttaagtgaatttttgccttcagcaaca 1949
Qy 2219 caatatggctcaggtcacttttttgaatgtcatttggcgccagat 2268
Dy 1950 caatatggctcaggtcacttttttgaatgtcatttggcgccagat 1999

RESULT 9
AAF21447
ID AAF21447 standard; DNA; 5036 BP.
XX
AC AAF21447;
XX DT
XX 14-MAR-2001 (first entry)
XX Human endothelin receptor A polynucleotide fragment #3014.
XX
KW Low adenosine antisease oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antialsthatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
XX Homo sapiens.
XX
XX WO200062736-A2.
XX
XX 26-OCT-2000.
XX
XX 24-MAR-2000; 2000WO-US08020.
XX
XX 06-APR-1999; 99US-0127958.
XX
XX (UYEC-) UNIV EAST CAROLINA.
XX PA
XX (NYCE/) NYCE J W.
XX
XX NYCE JW;
XX
XX WPI; 2000-679539/66.
XX
XX Low adenosine (A) content antisease oligonucleotides which do not
XX trigger adenosine receptors during metabolism, useful e.g. for treating
XX cancers and respiratory obstructions -
XX
XX Disclosure; Page 243-244; 1592pp; English.
XX
XX The present invention describes low adenosine (A) content antisease
XX oligonucleotides and compositions (I) comprising them. In the antisease
XX oligonucleotides the A is replaced by a 'universal' or alternative base.
XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
XX immunosuppressive, antialsthatic, hypotensive and cytostatic activities.
XX The antisease oligonucleotides and (I) can be used to down-regulate the
XX expression and/or activity of target polypeptides associated with
XX lung/respiratory disorders and malignancies, such as stimulating and
XX activating peptide factors and transmitters, transcription factors,
XX immunoglobulins and antibodies, antibody receptors, cytokines and
XX chemokines, endogenously produced specific and non-specific enzymes,
XX binding proteins, adhesion molecules and their receptors, cytokine and
XX chemokine receptors, adenosine receptors, bradykinin receptors, central
XX nervous system (CNS) and peripheral nervous and non-nervous system
XX receptors, CNS and peripheral nervous and non-nervous system peptide
XX transmitters, defensins, growth factors, vasoactive peptides and

CC receptors, binding proteins and malignancy associated proteins. The
CC antisease oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisease oligonucleotides used in the exemplification of
CC the present invention.
XX
XX

SQ Sequence 5036 BP; 1359 A; 1148 C; 1060 G; 1469 T; 0 other;

Query Match 44.3%; Score 1819.6; DB 21; Length 5036;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 421 aaaaagtgaagtgtaaaagcagcacagaagtgaataaagagatattctcaaatggcct 480
Dy 1310 aaaaagtgaagtgtaaaagcagcacagaagtgaataaagagatattctcaaatggcct 1369
Qy 481 caagatggaacccctttgctcagggcactcttttggctggcactggttgatgtgaat 540
Dy 1370 caagatggaacccctttgctcagggcactcttttggctggcactggttgatgtgaat 1429
Qy 541 cagtgataatcctcagagatagacacacaaataagcaaatcatgtggatgattccacac 600
Dy 1430 cagtgataatcctcagagatagacacacaaataagcaaatcatgtggatgattccacac 1489
Qy 601 tttctggtgacagagctcagcttctctgttaccactcatcaaccacataattgctct 660
Dy 1490 tttctggtgacagagctcagcttctctgttaccactcatcaaccacataattgctct 1549
Qy 661 acccagcaatggctcaatgacacactattgccacagcagactaaaaataactcagcttt 720
Dy 1550 acccagcaatggctcaatgacacactattgccacagcagactaaaaataactcagcttt 1609
Qy 721 caataacattaaacactgtgatatcttctactatttctcgtgggaatgggggaatgc 780
Dy 1610 caataacattaaacactgtgatatcttctactatttctcgtgggaatgggggaatgc 1669
Qy 781 aactctgctcaggtacattaccagaaacaaatgtatgaggaatggcccaacgcgctgat 840
Dy 1670 aactctgctcaggtacattaccagaaacaaatgtatgaggaatggcccaacgcgctgat 1729
Qy 841 agccagttctccttggagacacttatctatgtgtcattgatctccctataatgtatt 900
Dy 1730 agccagttctccttggagacacttatctatgtgtcattgatctccctataatgtatt 1789
Qy 901 taagctgctgctggcgctggccttttgatcacaaatgaactttggcgatattctttgaa 960
Dy 1790 taagctgctgctggcgctggccttttgatcacaaatgaactttggcgatattctttgaa 1849
Qy 961 gctgtcccttttgcagaagctcctcgtgggagacacgcctcccaacctcgcctct 1020
Dy 1850 gctgtcccttttgcagaagctcctcgtgggagacacgcctcccaacctcgcctct 1909
Qy 1021 tagtgtgacaggtacagagcagttgctcctcgtgagctggttcagggaattgggattcc 1080
Dy 1910 tagtgtgacaggtacagagcagttgctcctcgtgagctggttcagggaattgggattcc 1969
Qy 1081 tttgtaactgccaattgaaattgtctccatctggtacatctgctcttaccctggccatcc 1140
Dy 1970 tttgtaactgccaattgaaattgtctccatctggtacatctgctcttaccctggccatcc 2029
Qy 1141 tgaagcattggtcctcgtcattgggtacccctttgaatatagggtgaaacagcaataaacctg 1200
Dy 2030 tgaagcattggtcctcgtcattgggtacccctttgaatatagggtgaaacagcaataaacctg 2089

QY 1201 tatgtcaatgcacatcaaaatctatgagttctaccagaatgtaaagactggtggtc 1260
|||||
Db 2090 tatgtcaatgcacatcaaaatctatgagttctaccagaatgtaaagactggtggtc 2149
|||||
QY 1261 ctctgggtctctattctgtatgcccttggtgtagcactggatcttataccctcatgac 1320
|||||
Db 2150 ctctgggtctctattctgtatgcccttggtgtagcactggatcttataccctcatgac 2209
|||||
QY 1321 ttgtgagatgttgacacagaaggaatgcagcttgagaattgcctcagtgaaacatcttaa 1380
|||||
Db 2210 ttgtgagatgttgacacagaaggaatgcagcttgagaattgcctcagtgaaacatcttaa 2269
|||||
QY 1381 gcagcgtcgagaagtgagcaaacagtttctgctggttgtaattttgctctttgctg 1440
|||||
Db 2270 gcagcgtcgagaagtgagcaaacagtttctgctggttgtaattttgctctttgctg 2329
|||||
QY 1441 gttccctctctactaaagccgtatattgaagaaaactggtataacgaatggacaga 1500
|||||
Db 2330 gttccctctctactaaagccgtatattgaagaaaactggtataacgaatggacaga 2389
|||||
QY 1501 ccgattgtaattacttagtttcttactgctcatggtattacatcggtattaaacttggcaac 1560
|||||
Db 2390 ccgattgtaattacttagtttcttactgctcatggtattacatcggtattaaacttggcaac 2449
|||||
QY 1561 catgaattcaatgtataaaacccatagctctgtattttgtgacagaataatttaaaattg 1620
|||||
Db 2450 catgaattcaatgtataaaacccatagctctgtattttgtgacagaataatttaaaattg 2509
|||||
QY 1621 ttccagtcgatccctgctgctgtgtttaccagtcacaaaagtctgatgacctcggtccc 1680
|||||
Db 2510 ttccagtcgatccctgctgctgtgtttaccagtcacaaaagtctgatgacctcggtccc 2569
|||||
QY 1681 catgaacggaaacacatccagtggaagacacacacatcaaaacacacacacacacac 1740
|||||
Db 2570 catgaacggaaacacatccagtggaagacacacacacacacacacacacacacac 2629
|||||
QY 1741 gagcagccataaggacagatgaactgaaccccttagaagcactcctcggtactcccat 1800
|||||
Db 2630 gagcagccataaggacagatgaactgaaccccttagaagcactcctcggtactcccat 2689
|||||
QY 1801 aatcctctcgagaaaaaatcacaggcaactgtgactccgggaatctctctctctgac 1860
|||||
Db 2690 aatcctctcgagaaaaaatcacaggcaactgtgactccgggaatctctctctctgac 2749
|||||
QY 1861 ctctctcttaattcactccac 1919
|||||
Db 2750 ctctctcttaattcactccac 2809
|||||
QY 1920 actggtttatccacccac 1979
|||||
Db 2810 actggtttatccacccac 2869
|||||
QY 1980 attctgctgtgtattcagcactaaaaaatggtggagctggggaggaatgaagactgt 2039
|||||
Db 2870 attctgctgtgtattcagcactaaaaaatggtggagctggggaggaatgaagactgt 2929
|||||
QY 2040 taaatgaacacagaagatatttactacttttgcattgataaataagactttcaagtacatg 2099
|||||
Db 2930 taaatgaacacagaagatatttactacttttgcattgataaataagactttcaagtacatg 2989
|||||
QY 2100 gctagcttttatggcagttctggtgaatttcaatgggaactgggtcaccatgaacacttta 2159
|||||
Db 2990 gctagcttttatggcagttctggtgaatttcaatgggaactgggtcaccatgaacacttta 3049
|||||
QY 2160 gagattaacgacagaattttactacttttttaagtga-tttttgtccttcagccaaaca 2218
|||||
Db 3050 gagattaacgacagaattttactacttttttaagtgaattttttgtccttcagccaaaca 3109
|||||
QY 2219 caatatgggtcaggtcacttttattgaaatgtcatttgggtgcagat 2268
|||||
Db 3110 caatatgggtcaggtcacttttattgaaatgtcatttgggtgcagat 3159
|||||

RESULT 10

AAF21435

XX AAF21435 standard; DNA; 117609 BP.

AC AAF21435;

XX DT 14-MAR-2001 (first entry)

XX Human receptor-related antisense polynucleotide #3002.

XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

KW human; airway disorder; bronchoconstriction; lung inflammation;

KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;

KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;

KW respiratory obstruction; pulmonary obstruction; impeded respiration;

KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;

KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;

KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;

KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;

KW cancer; ss.

XX Homo sapiens.

OS

XX WO200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000WO-US08020.

XX 06-APR-1999; 99US-0127958.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

PI Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not
trigger adenosine receptors during metabolism, useful e.g. for treating
cancers and respiratory obstructions -

PS Disclosure; Page 17-47; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense
oligonucleotides and compositions (I) comprising them. In the antisense
oligonucleotides the A is replaced by a 'Universal' or alternative base.
XX (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
XX The antisense oligonucleotides and (I) can be used to down-regulate the
expression and or activity of target polypeptides associated with
lung/respiratory disorders and malignancies, such as stimulating and
activating peptide factors and transmitters, transcription factors,
immunoglobulins and antibodies, antibody receptors, cytokines and
chemokines, endogenously produced specific and non-specific enzymes,
binding proteins, adhesion molecules and their receptors, cytokine and
chemokine receptors, adenosine receptors, bradykinin receptors, central
nervous system (CNS) and peripheral nervous and non-nervous system
receptors, CNS and peripheral nervous and non-nervous system peptide
transmitters, defensins, growth factors, vasoactive peptides and
receptors, binding proteins and malignancy associated proteins. The
antisense oligonucleotides may be used in this way to treat disorders
including respiratory obstruction (especially pulmonary obstruction
and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
and/or surfactant hypoproduction which are associated with a disease or
condition selected from pulmonary vasoconstriction, inflammation,
allergies, asthma, impeded respiration, respiratory distress syndrome
(RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
pulmonary transplantation rejection, pulmonary infections, bronchitis,
and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
fragments and antisense oligonucleotides used in the exemplification of
the present invention.


```
XX SQ Sequence 117609 BP: 27530 A: 29942 C: 30320 G: 29441 T: 376 other:
Query Match 44.3%; Score 1819.6; DB 21; Length 117609;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

Qy 421 aaaaagtgaaggtgtaaagcagcacaaagtgcataaagagatatttctcctcaaaatttgctt 480
Db 87367 aaaaagtgaaggtgtaaagcagcacaaagtgcataaagagatatttctcctcaaaatttgctt 87426

Qy 481 caagatgaagaaccccttttgcctcagggaatcccttttggctggcactgggtgagtgataat 540
Db 87427 caagatgaagaaccccttttgcctcagggaatcccttttggctggcactgggtgagtgataat 87486

Qy 541 cagtataatcctgagagatcacagcacaaatctaagcaatcatgtggatgattccacac 600
Db 87487 cagtataatcctgagagatcacagcacaaatctaagcaatcatgtggatgattccacac 87546

Qy 601 ttttggccagagctcagcttctcttctgtttaccactcatcaaccactaatttggctct 660
Db 87547 ttttggccagagctcagcttctcttctgtttaccactcatcaaccactaatttggctct 87606

Qy 661 acccagcaatggctcaatgcacaactattgcccacagcagactaaaattacttcagcttt 720
Db 87607 acccagcaatggctcaatgcacaactattgcccacagcagactaaaattacttcagcttt 87666

Qy 721 caaatataactaacctgtgatatcttgaatatttcaatcttggaatgggtggaatgc 780
Db 87667 caaatataactaacctgtgatatcttgaatatttcaatcttggaatgggtggaatgc 87726

Qy 781 aactctgctcaggatcatattaccagaaacaaatgtatgaggaatggcccaacgcgctgat 840
Db 87727 aactctgctcaggatcatattaccagaaacaaatgtatgaggaatggcccaacgcgctgat 87786

Qy 841 agccagcttgcctctggagaccttatctatgttggtcattgatctccatcaaatgatt 900
Db 87787 agccagcttgcctctggagaccttatctatgttggtcattgatctccatcaaatgatt 87846

Qy 901 taagctgctgctggcgtggccttttgatcacaaatgaactttggcgtatttctttgcaa 960
Db 87847 taagctgctgctggcgtggccttttgatcacaaatgaactttggcgtatttctttgcaa 87906

Qy 961 gctgttccctttttgcagaagtcctcctgggtgggataccctcctcaacctctgcctct 1020
Db 87907 gctgttccctttttgcagaagtcctcctgggtgggataccctcctcaacctctgcctct .87966

Qy 1021 tagtgtgacaggtacagacagcttgcctcctggagtcgtgttcagggaattgggattcc 1080
Db 87967 tagtgtgacaggtacagacagcttgcctcctggagtcgtgttcagggaattgggattcc 88026

Qy 1081 tttggttaactgccattgaaattgtctccatctggatcctgtccctttatccctggccattcc 1140
Db 88027 tttggttaactgccattgaaattgtctccatctggatcctgtccctttatccctggccattcc 88086

Qy 1141 tgaagcgaattggctcgtcatggtaccctttgatatagggttgaaacagcataaaacctg 1200
Db 88087 tgaagcgaattggctcgtcatggtaccctttgatatagggttgaaacagcataaaacctg 88146

Qy 1201 tatgtcaatgccacacataaaattcatggattcttaccagaatgataaaggactggtgctt 1260
Db 88147 tatgtcaatgccacacataaaattcatggattcttaccagaatgataaaggactggtgctt 88206

Qy 1261 cttcgggttctatttctgtatgcccttgggtgctgcactgcgactcttctacacccctcatgac 1320
Db 88207 cttcgggttctatttctgtatgcccttgggtgctgcactgcgactcttctacacccctcatgac 88266

Qy 1321 ttgtgagatgttgaaacagaaggaatggcagcttgagattgcccctcagtgaaacatcttaa 1380
Db 88267 ttgtgagatgttgaaacagaaggaatggcagcttgagattgcccctcagtgaaacatcttaa 88326

Qy 1381 gcagcgtcgagaagtgccaaacacagtttcttctgcttgggtgtgtaatttttgcctcttgc 1440
```

RESULT 11

AAF20902

ID AAF20902 standard; DNA; 1310 BP.

XX AAF20902;

XX AAF20902;

DT 14-MAR-2001 (first entry)

XX

Human endothelin receptor A polynucleotide fragment #2469.

DE

XX

Low adenosine antisense oligonucleotide; phosphorothioate; allergy;

KW

human; airway disorder; bronchoconstriction; lung inflammation;

KW

surfactant depletion; respiratory; bronchodilator; antiinflammatory;

immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic; respiratory obstruction; pulmonary obstruction; impeded respiration; surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS; respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis; pulmonary hypertension; emphysema; pulmonary transplantation rejection; chronic obstructive pulmonary disease; pulmonary infection; bronchitis; cancer; ss.

Homo sapiens.

W0200062736-A2.

26-OCT-2000.

24-MAR-2000; 2000WO-US08020.

06-APR-1999; 99US-0127958.

(UYEC-) UNIV EAST CAROLINA.

(NYCE/) NYCE J W.

Nyce JW;

WPI; 2000-679539/66.

Low adenosine (A) content antisense oligonucleotides which do not trigger adenosine receptors during metabolism, useful e.g. for treating cancers and respiratory obstructions -

Disclosure; Page 244-245; 1592pp; English.

The present invention describes low adenosine (A) content antisense oligonucleotides and compositions (I) comprising them. In the antisense oligonucleotides the A is replaced by a 'Universal' or alternative base. (I) can have respiratory, bronchodilator, antiinflammatory, analgesic, immunosuppressive, antiasthmatic, hypotensive and cytostatic activities. The antisense oligonucleotides and (I) can be used to down-regulate the expression and/or activity of target polypeptides associated with lung/respiratory disorders and malignancies, such as stimulating and activating peptide factors and transmitters, transcription factors, immunoglobulins and antibodies, antibody receptors, cytokines and chemokines, endogenously produced specific and non-specific enzymes, binding proteins, adhesion molecules and their receptors, cytokine and chemokine receptors, adenosine receptors, bradykinin receptors, central nervous system (CNS) and peripheral nervous and non-nervous system receptors, CNS and peripheral nervous and non-nervous system peptide transmitters, defensins, growth factors, vasoactive peptides and receptors, binding proteins and malignancy associated proteins. The antisense oligonucleotides may be used in this way to treat disorders including respiratory obstruction (especially pulmonary obstruction and/or bronchoconstriction) and/or lung inflammation, allergies and/or surfactant hypoproduction which are associated with a disease or condition selected from pulmonary vasoconstriction, inflammation, allergies, asthma, impeded respiration, respiratory distress syndrome (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), pulmonary transplantation rejection, pulmonary infections, bronchitis, and/or cancer. AAF18434 to AAF21543 represent human polynucleotide fragments and antisense oligonucleotides used in the exemplification of the present invention.

Sequence 1310 BP; 337 A; 312 C; 284 G; 377 T; 0 other;

Query Match 31.6%; Score 1296.8; DB 21; Length 1310;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1301; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 481 caagatggaacccttgctcaggagatccttttgctgctggcactggttgatgtaat 540
|| |||||
Db 3 cacatggaacccttgctcaggagatccttttgctgctggcactggttgatgtaat 62
QY 541 cagtgataactcctgagatagacagcacaataactaagcaatcatgtggatgtttccaccac 600

Db 63 cagtgataactcctgagatagacagcacaataactaagcaatcatgttgatgtttccaccac 122
QY 601 ttttcgtgacagagcagctcctggttaccactcatcaaccacataattggtcct 660
Db 123 ttttcgtgacagagcagctcctggttaccactcatcaaccacataattggtcct 182
QY 661 acccagcaatggctcaatgacacactattgcccacagcagactaaaattacttcaagcttt 720
Db 183 acccagcaatggctcaatgacacactattgcccacagcagactaaaattacttcaagcttt 242
QY 721 caatacattaaacactgtgatacttctgtactattttctcgtgggaatggtgggaatgc 780
Db 243 caatacattaaacactgtgatacttctgtactattttctcgtgggaatggtgggaatgc 302
QY 781 aactctgctcagatcatttaccagaacaaaatgtatgaggaatggcccacacgcgtgat 840
Db 303 aactctgctcagatcatttaccagaacaaaatgtatgaggaatggcccacacgcgtgat 362
QY 841 agccagctctgccccttgagacattatctatgtgctcattgatctccctcatcaatgatt 900
Db 363 agccagctctgccccttgagacattatctatgtgctcattgatctccctcatcaatgatt 422
QY 901 taagctgctggctggcgtgcttcttggatcacaaatgactttggcgtattttttgcaa 960
Db 423 taagctgctggcgtggcgtgcttcttggatcacaaatgactttggcgtattttttgcaa 482
QY 961 gctgttccccttttgcagaagcctcctgggggataccgctcctcaacacctctgcctct 1020
Db 483 gctgttccccttttgcagaagcctcctgggggataccgctcctcaacacctctgcctct 542
QY 1021 tagtgtacaggtacacagcagcttgcctccctggagctgttccaggggaatgggattcc 1080
Db 543 tagtgtacaggtacacagcagcttgcctccctggagctgttccaggggaatgggattcc 602
QY 1081 ttgtgtaactgctcattgaaattgctccatctggatcctgctcttcttctcctggcattcc 1140
Db 603 ttgtgtaactgctcattgaaattgctccatctggatcctgctcttcttctcctggcattcc 662
QY 1141 tgaacgattggctcgtcattgtaacctttgaaataagggttgacacacataaaacctg 1200
Db 663 tgaacgattggctcgtcattgtaacctttgaaataagggttgacacacataaaacctg 722
QY 1201 tatgctcaatgccacatacaaaatcatgagttctaccacagatgtaaaagactggtgct 1260
Db 723 tatgctcaatgccacatacaaaatcatgagttctaccacagatgtaaaagactggtgct 782
QY 1261 ctccgggttctatttctgtatgccccttgggtgctcgtcgtctcttctacacctcatgac 1320
Db 783 ctccgggttctatttctgtatgccccttgggtgctcgtcgtctcttctacacctcatgac 842
QY 1321 ttgtgagatgttgaacagaaggaatggcagcttgagaattgcctcagtgaaacatttaa 1380
Db 843 ttgtgagatgttgaacagaaggaatggcagcttgagaattgcctcagtgaaacatttaa 902
QY 1381 gcagcgtcgaagtgccaaacacagtttctcgtggttgtaatttttgcctcttgcctg 1440
Db 903 gcagcgtcgaagtgccaaacacagtttctcgtggttgtaatttttgcctcttgcctg 962
QY 1441 gttccctcttcaacttaagccgtatatgtgaagaaaactgtgtataacgaataatggacaagaa 1500
Db 963 gttccctcttcaacttaagccgtatatgtgaagaaaactgtgtataacgaataatggacaagaa 1022
QY 1501 ccgagtgaattacttagtttcttactgctcatgattacacgttgatttaacttggcaac 1560
Db 1023 ccgagtgaattacttagtttcttactgctcatgattacacgttgatttaacttggcaac 1082
QY 1561 catgaattcatgataaaaccccatgctgtattttgtgacagaataatttaaaattg 1620
Db 1083 catgaattcatgataaaaccccatgctgtattttgtgacagaataatttaaaattg 1142
QY 1621 ttccagtcagctcctgctgctgtgttaccagtcacaaagtctgtatgacacctcggtccc 1680

Db 243 caaatattaaactgtgatattctgttactattttctatctgctgggaatggtgggaatgc 302
QY 781 aactctgctcaggatcatttaccagaacaaatgtatgaggaatggccccaacgcgctgat 840
Db 303 aactctgctcaggatcatttaccagaacaaatgtatgaggaatggccccaacgcgctgat 362
QY 841 agccagctctgccccttgagacacttatctatgtgtggttcaattgatactcactcaatgatt 900
Db 363 agccagctctgccccttgagacacttatctatgtgtggttcaattgatactcactcaatgatt 422
QY 901 taagctctgctggtggcgtgccccttttgatcaacaatgaacttggcttattttttgcaa 960
Db 423 taagctctgctggtggcgtgccccttttgatcaacaatgaacttggcttattttttgcaa 482
QY 961 gctgttcccctttttgagaagctcctggtgggatcacgctcctcaacccctgcgctct 1020
Db 483 gctgttcccctttttgagaagctcctggtgggatcacgctcctcaacccctgcgctct 542
QY 1021 tagtgttgacaggtacagagcagtgctcctcctggagtcgtgttccagggaattgggattcc 1080
Db 543 tagtgttgacaggtacagagcagtgctcctcctggagtcgtgttccagggaattgggattcc 602
QY 1081 ttgtgtaactgccattgaaattgtctccatctggtcctctggtccttcttctcctggtccattcc 1140
Db 603 ttgtgtaactgccattgaaattgtctccatctggtcctcctggtccttcttctcctggtccattcc 662
QY 1141 tgaagcgtattggtctgctcgtgacccctttgaaatagaggtggaacagcagataaaacctg 1200
Db 663 tgaagcgtattggtctgctcgtgacccctttgaaatagaggtggaacagcagataaaacctg 722
QY 1201 tagtctcaatgccacatcaaaaattcatggagttctaccagatgtataagagactggtggtct 1260
Db 723 tagtctcaatgccacatcaaaaattcatggagttctaccagatgtataagagactggtggtct 782
QY 1261 ctccggttctatttctgtatgcccctggtgtgacactgcgactcttctacacctctcagac 1320
Db 783 ctccggttctatttctgtatgcccctggtgtgacactgcgactcttctacacctctcagac 842
QY 1321 ttgtgagatgtgaacagaaggaatggcagcttgagaaatgcccctcagtgaaacatctaa 1380
Db 843 ttgtgagatgtgaacagaaggaatggcagcttgagaaatgcccctcagtgaaacatctaa 902
QY 1381 gcaagctcagaagtggaacaaacagtttctgctggtgtgtaatttttctctcttctgctg 1440
Db 903 gcaagctcagaagtggaacaaacagtttctgctggtgtgtaatttttctctcttctgctg 962
QY 1441 gttccctcttcaacttaagcgtatattgaagaactgtataacgaagaatggacaagaa 1500
Db 963 gttccctcttcaacttaagcgtatattgaagaactgtataacgaagaatggacaagaa 1022
QY 1501 ccgagtgaattacttagttcttactgctcagtgattacatcgttatttaacttggcgaac 1560
Db 1023 ccgagtgaattacttagttcttactgctcagtgattacatcgttatttaacttggcgaac 1082
QY 1561 catgaattcatgtataaacccccatagctctgtatttttggagcaagaattttaaattg 1620
Db 1083 catgaattcatgtataaacccccatagctctgtatttttggagcaagaattttaaattg 1142
QY 1621 ttccagtcagctcctgctgctgctgttaccagtcacaaagctcagctcctgctcctc 1680
Db 1143 ttccagtcagctcctgctgctgctgttaccagtcacaaagctcagctcctgctcctc 1202
QY 1681 catgaacggaacaaagcatccagtggaagaccacgatcaaaaacacacacacacagcgg 1740
Db 1203 catgaacggaacaaagcatccagtggaagaccacgatcaaaaacacacacacacagcgg 1262
QY 1741 gagcagccataaggacagcatgaactgaccaccccttagaagcactct 1788
Db 1263 gagcagccataaggacagcatgaactgaccaccccttagaagcactct 1310

RESULT 15
AAF21442

ID XX AAF21442 standard; DNA; 146981 BP.
XX AC AAF21442;
XX XX
DT 14-MAR-2001 (first entry)
XX
XX Human ELAM-1 polynucleotide fragment #3009.
DE
XX
KW Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
KW human; airway disorder; bronchoconstriction; lung inflammation;
KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;
KW respiratory obstruction; pulmonary obstruction; impeded respiration;
KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
KW cancer; ss.
XX
XX Homo sapiens.
OS
XX
XX WO2000062736-A2.
PN
XX
XX 26-OCT-2000.
PD
XX
XX 24-MAR-2000; 2000WO-US08020.
PF
XX
XX 06-APR-1999; 99US-0127958.
PR
XX
XX (UYEC-) UNIV EAST CAROLINA.
PA
XX (NYCE/) NYCE J W.
PI
XX Nyce JW;
XX
XX WPI; 2000-679539/66.
DR
XX
XX Low adenosine (A) content antisense oligonucleotides which do not
PT trigger adenosine receptors during metabolism, useful e.g. for treating
PT cancers and respiratory obstructions -
XX
XX Disclosure; Page 146-177; 1592pp; English.
PS
XX
XX The present invention describes low adenosine (A) content antisense
CC oligonucleotides and compositions (I) comprising them. In the antisense
CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.
CC The antisense oligonucleotides and (I) can be used to down-regulate the
CC expression and/or activity of target polypeptides associated with
CC lung/respiratory disorders and malignancies, such as stimulating and
CC activating peptide factors and transmitters, transcription factors,
CC immunoglobulins and antibodies, antibody receptors, cytokines and
CC chemokines, endogenously produced specific and non-specific enzymes,
CC binding proteins, adhesion molecules and their receptors, cytokine and
CC chemokine receptors, adenosine receptors, bradykinin receptors, central
CC nervous system (CNS) and peripheral nervous and non-nervous system
CC receptors, CNS and peripheral nervous and non-nervous system peptide
CC transmitters, defensins, growth factors, vasoactive peptides and
CC receptors, binding proteins and malignancy associated proteins. The
CC antisense oligonucleotides may be used in this way to treat disorders
CC including respiratory obstruction (especially pulmonary obstruction
CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
CC and/or surfactant hypoproduction which are associated with a disease or
CC condition selected from pulmonary vasoconstriction, inflammation,
CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
XX Sequence 146981 BP; 44332 A; 30090 C; 28076 G; 44464 T; 19 other;

Query Match	31.5%	Score	1296.8	DB	21	Length	146981
Best Local Similarity	99.5%	Pred. No.	0				
Matches	1301	Conservative	0	Mismatches	7	Indels	0
Gaps	0						

QY	481	caagatgaaaccccttgccctcaggagcatccctttgctgcgcactggttgatgtgtaatt	540
DB	145674	ccccatggaaccccttgccctcaggagcatccctttgctgcgcactggttgatgtgtaatt	145733
QY	541	cagtataactctgagagatacagcacaaaaattcaagcaaatcatgtggatgatttcacccac	600
DB	145734	cagtataactctgagagatacagcacaaaaattcaagcaaatcatgtggatgatttcacccac	145793
QY	601	ttttcgtggcaagagctcagctccctgggttacacatcatcaaacccactaatttggctcct	660
DB	145794	ttttcgtggcaagagctcagctccctgggttacacatcatcaaacccactaatttggctcct	145853
QY	661	accagcaatggctcaatgcacaaactattgccacagcagactaaaaattacttcagcttt	720
DB	145854	accagcaatggctcaatgcacaaactattgccacagcagactaaaaattacttcagcttt	145913
QY	721	caatacaattaaactgtgatattcttgactattttctatctgtgggaattggtgggaatgc	780
DB	145914	caatacaattaaactgtgatattcttgactattttctatctgtgggaattggtgggaatgc	145973
QY	781	aactctgctcaggatcatttaccagacaacaaatgataggaatggccccaacgcgctgat	840
DB	145974	aactctgctcaggatcatttaccagacaacaaatgataggaatggccccaacgcgctgat	146033
QY	841	agccagcttgcccttgagacatttatctatgttggtcattgatctccctatcaattgtatt	900
DB	146034	agccagcttgcccttgagacatttatctatgttggtcattgatctccctatcaattgtatt	146093
QY	901	taactgctggctggcgctggcccttttgatcacaatgaactttggcgtattttctttgcaa	960
DB	146094	taactgctggctggcgctggcccttttgatcacaatgaactttggcgtattttctttgcaa	146153
QY	961	gctgtccccccttttgagaagatccctgggtggggatcacccctcccaacctctgcgcct	1020
DB	146154	gctgtccccccttttgagaagatccctgggtggggatcacccctcccaacctctgcgcct	146213
QY	1021	tagtgtgaacaggtacagagcagttgtccctcctggagtcggttcaagggaattgggattcc	1080
DB	146214	tagtgtgaacaggtacagagcagttgtccctcctggagtcggttcaagggaattgggattcc	146273
QY	1081	tttggttaactgccattgaaattgtctccatctggatctgtctcttattcctggccattcc	1140
DB	146274	tttggttaactgccattgaaattgtctccatctggatctgtctcttattcctggccattcc	146333
QY	1141	tgaagcgattggcttcgtctcatggttacccttttgaaatagatggggtgaaacagcataaacctg	1200
DB	146334	tgaagcgattggcttcgtctcatggttacccttttgaaatagatggggtgaaacagcataaacctg	146393
QY	1201	tatgtcaatgccacatacaaaattcatggagttctaccaagaatgataaaggactgctggct	1260
DB	146394	tatgtcaatgccacatacaaaattcatggagttctaccaagaatgataaaggactgctggct	146453
QY	1261	cttcgggtttctatttctgtatgcccttggttgcaactcgatcttctacacctcatgac	1320
DB	146454	cttcgggtttctatttctgtatgcccttggttgcaactcgatcttctacacctcatgac	146513
QY	1321	ttgtgagatgttgaaacagaaggaatggcagcttgagaattgccctcagtgaaacattctaa	1380
DB	146514	ttgtgagatgttgaaacagaaggaatggcagcttgagaattgccctcagtgaaacattctaa	146573
QY	1381	gcagcgtcgaagaagtggcacaacacagttttctgcttggttgtaatttttgcctttgctg	1440
DB	146574	gcagcgtcgaagaagtggcacaacacagttttctgcttggttgtaatttttgcctttgctg	146633
QY	1441	gttccctctcaacttaagccgttatattgaagaaaaactgtgtatacagaatggacaagaa	1500
DB	146634	gttccctctcaacttaagccgttatattgaagaaaaactgtgtatacagaatggacaagaa	146693

Search completed: October 1, 2002, 01:01:38
Job time: 28537 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2002, 17:02:56 ; Search time 14667 Seconds
(without alignments)
5856.935 Million cell updates/sec

Title: US-09-931-157-1
Perfect score: 4105
Sequence: 1 gaattcgccgcctcttg.....taaaagtttacagaacctt 4105

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.htg.*
- 3: gb.in.*
- 4: gb.om.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.or.*
- 22: em.ov.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htgo.inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Score	Match Length	DB ID	Description
-----	-----	-----	-----	-----

1	4105	100.0	4105	6	AR177879
2	4105	100.0	4105	6	E07649
3	4105	100.0	4105	9	HS21R
4	4103	100.0	4105	9	S57498
5	4101	99.9	4101	9	HUMETNLR
6	3183.2	77.5	3305	9	S45956
7	2609.6	63.6	2705	9	BC022511
8	2470	60.2	2595	9	HUMETAR8
9	2409.2	58.7	164920	9	AC093908
10	2292	55.8	2337	11	G06463
11	1819.6	44.3	1868	9	S63938
12	1786.6	43.5	3216	4	BTBETREC
13	1560.2	38.0	1661	9	S67127
14	1359	33.1	1359	9	S81539
15	1296.8	31.6	1310	9	HUMEDNRA
16	1280.8	31.2	1284	6	AX280871
17	1122	27.3	1374	4	S80652
18	1103.2	26.9	1284	4	AF416703
19	1088.2	26.5	2696	10	BC008277
20	1048.8	25.5	1180	4	AF311974
21	1021	24.9	1021	11	G10643
22	997	24.3	1436	10	RATENDOR
23	951	23.2	1160	9	S81542
24	765.8	18.7	2344	5	AF040634
25	695	16.9	1032	9	S81545
26	632.8	15.4	1650	5	XL006633
27	624.6	15.2	810	10	AF039892
28	514	12.5	530	9	MFU20577
29	492.4	12.0	531	9	HUMETAR2
30	487.4	11.9	1435	9	S5554752
31	487.4	11.9	6264	2	AC110065
32	425	10.4	1308	5	CCEDNRB2
33	411.6	10.0	1724	5	AB045356
34	411	10.0	1314	4	AF276427
35	407.8	9.9	1311	4	AF038900
36	407.8	9.9	1452	4	AF034530
37	403.8	9.8	1892	10	RNETBREC
38	403.8	9.8	1965	6	E03623
39	403.8	9.8	2018	10	S65355
40	403	9.8	1326	4	AF245469
41	403	9.8	1669	4	AF019072
42	401.4	9.8	2026	4	BOVEETBR
43	401	9.8	1309	9	HUMETAR1
44	398.4	9.7	1041	5	CCEDNRB
45	396.6	9.7	1551	6	E05930

ALIGNMENTS

RESULT 1	AR177879	Sequence 1 from patent US 6313276.	4105 bp	DNA	linear	PAT 17-DEC-2001
LOCUS	AR177879	Sequence 1 from patent US 6313276.				
DEFINITION	AR177879					
ACCESSION	AR177879					
VERSION	AR177879.1	GI:17920234				
KEYWORDS	Unknown.					
SOURCE	Unknown.					
ORGANISM	Unclassified.					
REFERENCE	1 (bases 1 to 4105)					
AUTHORS	Imura, H., Nakao, K. and Nakanishi, S.					
TITLE	Human endothelin receptor					
JOURNAL	Patent: US 6313276-A 1 06-NOV-2001;					
FEATURES	Location/Qualifiers					
source	1..4105					
BASE COUNT	1138 a 859 c 845 g 1263 t					
ORIGIN						
Query Match	100.0%;	Score 4105;	DB 6;	Length 4105;		
Best Local Similarity	100.0%;	Pred. No. 0;				

Matches 4105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 gaattcgcgcgcgcctcttgcgtccagagtggaagtgctggagcctttgggag 60
Db 1 GAAATTCGGGGCGGCTCTGCGGTCCAGAGTGAAGGTCTGGAGCTTTGGGAG 60
Qy 61 agacgggagacacactggagcgctgttctccggagtttctttcttcttgcgcgcct 120
Db 61 AGACGGGAGGACACACTGGAGCGGTGTCCTCCGGAGTTTCTTTTCGTGCGAGCCCT 120
Qy 121 cgcgcgcgtacagctatcccgctggctgacgatttggaagcggtggagagctt 180
Db 121 CGCGCGCGGTACAGTCAATCCGCTGGTCTGACGATTGTGGAGAGCGGTGGAGGCTT 180
Qy 181 catcaatccaccgcgcgtgcgcggggtattgggtccagcgacacctcccccggagaa 240
Db 181 CATCAATCCACCGCGTGCCTCCGGGATTTGGGTCCAGCGACACCTCCCGGGAGAA 240
Qy 241 gcagtcccaggaagtcttctgaagccggggaagctgtgcagccgaagcgcgcgcgc 300
Db 241 GCAGTGGCCAGGAAGTTTCTGAAGCCGGGGAAGCTGTGCAGCCGAAGCGCGCGCGGC 300
Qy 301 cggagccgggacacccgcacctccgcgcacccaccctcgctttctcgcgcttcc 360
Db 301 CGGAGCCGGGACACCGGCCACCTCCCGGCCACCCACCTCGCTTTCTCGGGCTTCCTC 360
Qy 361 tggccagcgccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
Db 361 TGGCCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
Qy 421 aaaaagtgaaggtgtaaagcagcaccaagtgcaataagagatatcttccaaattgct 480
Db 421 AAAAAGTGAAGGTGTAAAGCAGCACAAAGTGCATAAGAGATATTTCTCAAAATTGCT 480
Qy 481 caagatggaacccttgctccaggcacccttttgcgcgcgcgcgcgcgcgcgcgcgcgc 540
Db 481 CAAGATGGAACCCCTTGGCTCAGGGCATCTTTTGGCTGGCAGCTGGTGGATGTGTAAT 540
Qy 541 cagtgaataactcgtgagatgacagcaaaatacgaacatcatgtggatgtttccacc 600
Db 541 CAGTGAATAACTCCTGAGAGATACAGCACAAATCAAGCAATCATGTGGATGATTTCAACC 600
Qy 601 ttttcgtggcacagctcagcttccctgtttaccactcatcaaccacataatttgct 660
Db 601 TTTTCGTGGCACAGAGCTCAGCTTCCCTGGTTACCACCTCATCAACCCACTAATTTGGTCT 660
Qy 661 acccagcaatggctcaatgcacaactattgccacagcagactaaaattacttcagcttt 720
Db 661 ACCCAGCAATGGCTCAATGCACAATATTTGCCACAGCAGCACTAAATTTACTTCAGCTTT 720
Qy 721 caaatataactcgtgatctgttactattttcatctgtggaaatgggggaaatgc 780
Db 721 CAATATACATTAACACTGTGATATCTGTACTATTTTATCTCATCTGGGAATGGTGGGAATGC 780
Qy 781 aactctctcaggaatcattaccagaacaaatgtagggaatggcccccaacgcgtgat 840
Db 781 AACTCTCTCAGGATCATTTACCAGAACAAATGTATGAGGAATGGCCCCAACCGCGCTCAT 840
Qy 841 agccagcttgccttgagacattatctatgttggtcattgatctcccatcaatgtatt 900
Db 841 AGCCAGTCTTGGCCCTTGAGACCTTATCTATGTGGTCAATGATCTCCCTATCAATGTATT 900
Qy 901 taagctgctggctggcgctggccttttgatcacaaatgaactttggcgctatttcttgc 960
Db 901 TAAGCTCTGGCTGGCGCTGGCCCTTTTGATCACAAATGACTTTGGCGTATTTCTTTGCAA 960
Qy 961 gctgttccctttttgcagaagctcctcgtggggtatcacgcgtccacacctcgcgtct 1020
Db 961 GCTGTTCCCTTTTTCAGAAGCTCTCGGTGGGATCACCGTCCCTCAACCTCTCGGCTCT 1020
Qy 1021 tagtgttgacaggtacagacagctgctcctcgtgagctgctgttcagggaaatgggattcc 1080
Db 1021 TAGTGTGACAGGTACAGACAGTGTGCTCTCTGGAGTCTGTTTCAGGGAATTTGGGATTCC 1080
```

```
Qy 1081 tttggtaactgcoattgaaattgtctccatctcgtgactcttcttcttcttcttcttcc 1140
Db 1081 TTTGGTAACTGCCATTGAAATGTCTCCATCTGGATCTGCTCTTATCTCTGGCAATTC 1140
Qy 1141 tgaagcgattggcttgcgttcgtgacgcttacccttgaataataggggtgaacagacataaaacctg 1200
Db 1141 TGAAGCGATTGGCTTCGTTCATGCTACCTTTGAAATATAGGGGTGAACAGCATAAACCTG 1200
Qy 1201 tatgctcaatggcacatcaaaattcatggagttctccaaagatgtaaaagcagctggct 1260
Db 1201 TATGCTCAATGCCACATCAAAATTCATGAGTCTACCAAGATGTAAAGGACTTGGTGCT 1260
Qy 1261 cttcggggttctatttctgtatgcccttggctgacgtgcgtcttctacacctcatgac 1320
Db 1261 CTTCCGGTCTATTCTGTATGCCCTTGTGTGCTACCTGGATCTTCTACACCTCATGAC 1320
Qy 1321 ttgtgagatgttgaacagagaagaaatggcagcttgagaattgccctcagtgaaacatttaa 1380
Db 1321 TTGTGAGATGTGTAACAGAGGAATGGCAGCTTGAGAATTTGCCCTCAGTGAACATCTTAA 1380
Qy 1381 gcagcgtcgagagtggaagaaacagtttctgcttgggttgtaatttttgccttctgctg 1440
Db 1381 GCAGCGTCGAGAGTGGCAAAACAGTTTCTGCTTGGTGTGTAATTTTGTCTCTTCTGCTG 1440
Qy 1441 gttccctcttcaactaaagcgcgtatattgaagaaacctgtataacgaaatgacaagaa 1500
Db 1441 GTTCCCTCTTCACTTAAGCCGTATATTGAAGAAACTGTATTAACGAAATGGACAAAGAA 1500
Qy 1501 ccgagtgaattacttagtttcttctgctcgtgctgtttaccagtcacaaagctctgacccctgcgc 1560
Db 1501 CCGATGTGAATTTACTTAGTTTCTTACTGCTCATGGATTACATCGGTATTAACTTGGCAAC 1560
Qy 1561 catgaattcatatataaccctcagctcgtgtatttttgagcaagaataattttaaattg 1620
Db 1561 CATGAATTCATATATAACCCCATAGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTG 1620
Qy 1621 tttccagtcgctcgtcgtcgtgtgtttaccagtcacaaagctctgacccctgcgc 1680
Db 1621 TTTCCAGTCATGCCCTCTGCTGCTGTGTGTACCAGTCCAAAGTCTCATGACCTCGGTCCC 1680
Qy 1681 catgaacgaaacagcatccagtggaagaaacacacgatacaaaacacacacacagaccg 1740
Db 1681 CATGAACGGAACAGCATCCAGTGAAGAACACCATCATGATCAAAACACCAACACAGACCG 1740
Qy 1741 gagcagcataagggacagcatgaactgaccccttagaagcactcctcggtactcccat 1800
Db 1741 GAGCAGCCATAGGACAGCATGAATGACCCCTTTAGAACGACTCCTCGGTACTCCCAT 1800
Qy 1801 aatcctctcggagaaaaatacacaggcaactgtgactcccggaatctcttctctgac 1860
Db 1801 AATCCTCTCGGAGAAAAAATCACAGCAACTGTGACTCCGGGAATCTCTCTCTGATC 1860
Qy 1861 cttctctcttaattcaactccacaccccaagaagaatgctttcccaaaacgcgaagtga 1920
Db 1861 CTTCTTCTCTTAATTCACCTCCACACCAAGAAGAATGCTTTCCAAAACCCGAAGGTAGA 1920
Qy 1921 ctggtttatccaccacaaacatctcgaatcgtactcttcttcttcttcttcttcttct 1980
Db 1921 CTGTTTATCCACCACCAACATCTACGAATCTGTTTCTTTTAAATGATTAATTTACATA 1980
Qy 1981 tctgcgtgttcttctcagcactaaaaatggtggagctggggagagatgaagactgtt 2040
Db 1981 TTCGCGTGTGTATTACGACTAAAAAATGGTGGGAGCTGGGGGAGATGAAGACTGTT 2040
Qy 2041 aaatgaaccacgaaggaattatttactacttttgcataagaataagactttcaagtaatgg 2100
Db 2041 AAATGAACACCAAGAGGATATTACTACTTTTGCATGAAAAATAGAGCTTTCAAGTACATGG 2100
Qy 2101 ctagcttttatggcagttctcgtgaaatgttcaatgggaactggtccacatgaaactttag 2160
Db 2101 CTAGCTTTTATGGCAGTTCTGTTGAATGTTCAATGGGAACCTGTCACCATGAACCTTAG 2160
```


TITLE HUMAN ENDOTHELIN RECEPTOR
JOURNAL Patent: JP 1994157595-A 1 03-JUN-1994;
SHIONOGI & CO LTD
COMMENT OS Homo sapiens (human)
PN JP 1994157595-A/1
PD 03-JUN-1994
PF 12-JUL-1991 JP 1991172828
PI IMURA HIROO, NAKAO ICHIKAZU, NAKANISHI SHIGETADA PC
C07K13/00, C12N5/10, C12N15/12, C12P21/02, (C12N5/10, C12R1:91), PC
C12P21/02,
PC C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No; Location/Qualifiers
FH Key
FH source 1. 4105
/organism="Homo sapiens"
/tissue_type="placenta"
/clone="phETIR"
1. 484
485..1768
/product="endothelin receptor, ETA-receptor" FT
sig_peptide 485..544
/product="signal peptide of endothelin FT
receptor,
ETA-receptor"
FT mat_peptide 545..1765
FT 3'UTR 1769..4105
FT polyA_signal 4084..4089.
Location/Qualifiers
1. 4105
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 1138 a 859 c 845 g 1263 t
ORIGIN

Query Match 100.0%; Score 4105; DB 6; Length 4105;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 gaattcgccgcgccttcggtccagagtggaactgaaggtctggagcttgagg 60
DB 1 GAATTCGGCGGCCCTCTTTCGGGTCCAGAGTGAGGAGGCTGGAGCTTTGGGAGG 60

QY 61 agacggggaggacagactggagcggtgttcctccgagtttctttctgctgcgagccct 120
DB 61 ACACGGGAGGACAGACTGGAGCGGTTCCTCCGAGTTTCTTTTCGTGGAGCCCT 120

QY 121 cgcgcgcgtacagtcattccgcgtggtctgacgattgtggagagcggtgagagcgtt 180
DB 121 CGCGCGCGGCTACAGTCATCCGCGTGTCTGACGATTGTGGAGAGCGGTGGAGAGCCT 180

QY 181 catcatccaccgcgcgtctccgggagattgggtcccgagacacctccccgggagaa 240
DB 181 CATCATCCACCCGGTCTGCGCGGGGATTGGGGTCCAGCGACACCTCCCGGGAGAA 240

QY 241 cagtgccccaggaattttctgaagccggggaagctgtgcagccgaagccgcgcgcgcgc 300
DB 241 GCAGTGCACGAGAGTTTCTGAAGCCGGGGAAGCTGTGCAGCCGAAGCCGCGCGCGC 300

QY 301 cggagccggggacacgcgcacctccgcgcaccaccacctcgcttctccggttcctc 360
DB 301 CGGAGCCGGGACACCGGCCACCTCCGCGCACCCACACCTCGCTTCTCCGGCTTCCTC 360

QY 361 tggccagggcgccgcgcggagccgcagctgtctgcgcagccgcagctccacgggaa 420
DB 361 TGGCCAGGCGCGCGGACCGGACGCTGTCTGCGCACGCGGAGCTCCACGGTGA 420

QY 421 aaaaagtgaagtgtaaaagcagcaagtgcaataagagatatattcctcaatttgct 480

DB 421 AAAAGTGAAGGTCTAAAGCAGCACAGTCAATAGAGATATTTCTCAAAATTTGCT 480
QY 481 caagatggaacccctttgctcaggggcatccttttggctgcaactggttgatgtgta 540
DB 481 CAAGATGGAACCCCTTTGGCTCAGGGCATCTCTTTGGCTGCACTGGTGGATGTAT 540
QY 541 cagtgaatacctcagagatcacagacaaactaagcaatcatgtggatgtattcaccac 600
DB 541 CAGTGATAATCCCTGAGAGATACAGCAAAATCTAAGCAATCATGTGGATGATTTCCAC 600
QY 601 tttctggtggcacagagctcagcttcctgggtttaccactcatcaacccactaatttgg 660
DB 601 TTTTCTGGTGACAGAGCTCAGTTCCTGGTTTACCACATCATCAACCCACTAATTTGGTCT 660
QY 661 acccagcaatggctcaatgcacaactattgccacagcagactaaataacttaagcttt 720
DB 661 ACCCAGCAATGGCTCAATGCACAACTATTGCCACAGCAGACTAAAATTTACTTCACT 720
QY 721 caaatcattaaactgtgatcttctactatttctcgtgggaatggtgggaatgc 780
DB 721 CAATACATTTAACACTGTGATATCTTGTACTATTTTTCATCTGGGATGTTGGGAATGC 780
QY 781 aactctgctcagatcatattaccagaaactgtatgaggaatggtgggaatggtgggaatgc 840
DB 781 AACTCTGCTCAGGATCATTTACCAGAACTATGATGAGGAATGCCCCACAGCGCTGAT 840
QY 841 agcagctctgcccctggagaccttatctatgtgtggtcattgatctcctatcaatgtatt 900
DB 841 AGCAGTCTTTGCCCTTGGAGACCTTATCTATGTGTGCTGATTCCTCCATCAATGTAT 900
QY 901 taagctgctgctggcgctggtctttgatcacaaactgacattggtgctattcttctgcaa 960
DB 901 TAAGCTGCTGGCTGGCGCTGGCCCTTTTGATCAAAATGACCTTTGGCGTATTTCTTTG 960
QY 961 gctgttccccctttttgcagaagctcctcgtggggatcacccgctcctcaacctctgcgct 1020
DB 961 GCTGTTCCCCCTTTTGCAGAAAGTCTCGTGGGGATCACCGTCTCAACCTCTCGCTCT 1020
QY 1021 tagtggacaggtacagagcagctgctcctcctggagtcgtgttcagggaatggtgattcc 1080
DB 1021 TAGTGTGACAGGTACAGACAGCTGCTCTCTGGAGTGTGTTCAGGGAATTTGGGATTC 1080
QY 1081 ttggtaactgcctatgaaattgtctccatctgacatcctctcttctatctcctgcatcc 1140
DB 1081 TTTGGTAACGTGCTTGAATTTGTCTCCATCTGATCTGTCTTTATCTCTGGCCATTC 1140
QY 1141 tgaagcgattggcttcgtcatggttaccctttgaaataagggtgaaacagacataaaactg 1200
DB 1141 TGAAGCGATTGGCTTCGTCTATGTTGACCTTTGAATATAGGGGTGAACAGCATAAACCTG 1200
QY 1201 tatgtcaatgcacatcaaaaactcattgaggttctaccaagatgtaagagctggtgct 1260
DB 1201 TATGCTCAATGCCACATCAAAATTCATGGAGTCTTACAAAGATGTAAGGACTGGTGGCT 1260
QY 1261 ctccggttctattctgtatgctccttgggtgtgcactgcgactcttctacacctcatgac 1320
DB 1261 CTTCCGGGTCTATTTCTGTATGCTTGGTGTGCACGTGCGATCTTACACCCCTCATGAC 1320
QY 1321 ttgtgagattgtgaacagaggaatggcagcttgagaattgccctcagtgaaacatcttaa 1380
DB 1321 TTGTGAGATTGTGAACAGAGGAATGGCAGCTTGAGAAATTTGCCCTCAGTGAACATCTAA 1380
QY 1381 gacgctcagagaagtggaacaaacagtttctcgtggttctgtaatttttgcctttgctg 1440
DB 1381 GCAGCGTCGAGAAGTGGCAAAAACAGTTTCTGCTGGTGTGTAATTTTGTCTTTGCTG 1440
QY 1441 gttccctctcacttaagcgttatattgaagaaactgtataacaaatgacaaagaa 1500
DB 1441 GTTCCCTCTTCACTTAAGCGGTATATGAAGAAACCTGTGTATAACGAAATGGACAAGAA 1500
QY 1501 ccgattgtaattacttagttcttactgctcattggtattacatcggtatttaacttggcaac 1560

Db	1501	CCGATGTGAATTACTTAGTTTCTTACTGCTCATGGATTACATCGGTATTAACTTGGCAAC	1560
Qy	1561	catgaattcatgtataaacccccatagctctgtattttgtgagcaaaaaatttaaaaattg	1620
Db	1561	CATGAATTCATGTATAAACCCCATAGCTCTGTATTTGTGAGCAAGAAATTTAAAAATTG	1620
Qy	1621	tttccagtcagctctgtctgtctgtctgttaccagtcctcaaaagctctgatgacctcggtccc	1680
Db	1621	TTTCCAGTCATGCTCTGCTGCTGCTGTGTACCAGTCCAAAGTCTGATGACCTCGGTCCC	1680
Qy	1681	catgaacggaacaagcatccagctggaggaaccacgcgtacaaaaacacacacacagaccg	1740
Db	1681	CATGAACGGGAACAGCATCCAGTGGAGAACACGATCAAAACACCAACACACAGACCG	1740
Qy	1741	gagcagccataaagacagcatgaactgaccaccccttagaagcactcctcggtactcccat	1800
Db	1741	GAGCAGCCATAAGGACAGCATGAACTCACCACCCCTTAGAAGCACCTCCTCGGTACTCCCAT	1800
Qy	1801	aatcctctcgagaaaaaatcaaggaactgactccgggaactcctctctctgctgac	1860
Db	1801	AATCCCTCTCGGAGAAAAATCAAGGCAACTGTGACTCCGGGAATCTCTCTCTGATC	1860
Qy	1861	cttctctcttaattcactccacacaccccaagaagaatgctttccaaaaccgcaaggtaga	1920
Db	1861	CTTCTTCTTAATTACTCCACACCCACAGAGAANAATGCTTCCAAAACCGCAAGGTAGA	1920
Qy	1921	ctggtttatccaccacacatctcgaatcgtaacttctttaaattgatctaatttacata	1980
Db	1921	CTGGTTTATCCACCCCAACATCTAGCAATCGTACTCTTTAATTGATCTAAATTTACATA	1980
Qy	1981	ttctgcgtgtgtattcagcactaaaaatggtggagctggggaggaatgaagactggt	2040
Db	1981	TTCTGCGTGTGTATTTCAGCACCTAAAAAATGGTGGGAGCTGGGGGAGAAATGAAGACTGTT	2040
Qy	2041	aaatgaaccagagaagatatctactcttttgcatgaataagagctttcaagtcatacg	2100
Db	2041	AAATGAACACAGAGGATATTTACTACTTTTGCATGAANAATAGAGCTTTCAAGTACATGG	2100
Qy	2101	ctagcttttatggcagctctggtgaatgttcaatgggaactggccaccatgaacctttag	2160
Db	2101	CTAGCTTTTATGGCAGTCTGTGGTGAATGTTCAATGGGAACCTGCTACCATGAACCTTTAG	2160
Qy	2161	agattaaacacagaagatttctactctttttaaagtgatgttttgccttcagccaacaca	2220
Db	2161	AGATTAAACGACAAGATTTTCTACTTTTAAAGTGATTTTGTCTCTCAGCCAAAACACA	2220
Qy	2221	atatgggctcaggtcacttttatgtgaatgtcatttggcgagctatttttcaactgca	2280
Db	2221	ATATGGGCTCAGGTCACTTTTATTTGAATGTCAATTTGGTGGCCAGTATTTTAACTGCA	2280
Qy	2281	taatgacctaacatgattattgaaacttatttacacatagtttgaaaaaaaagaacaaa	2340
Db	2281	TAATAGCCTAAACATGATTTATTTGAACCTATTACACATAGTTTGAAAAAAAAAGACAA	2340
Qy	2341	aatagttatcaggtgagcaattagattagatttttccacgcactatttatttttttaa	2400
Db	2341	AATAGTATTTCAGGTGAGCAATTAGATTAGATTATTTCCACGCTCACTATTATTTTAA	2400
Qy	2401	acacaaattctaaagctacaacaaatactacaggcccttaaacagcagctctgatgacaca	2460
Db	2401	ACACAATTTCTAAGCTACAACNAATACTACAGGCCCTTAAGACAGCTCTGATGACACA	2460
Qy	2461	tttggcagtttaatatagatgttactcaagaattttttaagaactgatttttttttaa	2520
Db	2461	TTTGGCAGTTTAATAGATGTTACTCAAGAAATTTTAAAGAACTGTATTTTATTTTAA	2520
Qy	2521	atgggttttattacaagggaaccttgaacatgttttgtatgtttaaatccaagaatgac	2580
Db	2521	ATGGGTGTTTATACAAAGGACCTTGAACATGTTTTGTATGTATAATTCAAAAGTAATGC	2580
Qy	2581	ttcaatcacagatgttcttttcacaaagttcaatactgttttttcagtgaaatttgtatga	2640
Db	2581	TTCAATCACATAGTCTTTTTCACAAGTTCATACTGTGTTTTTTCATGTAATTTTGTATGA	2640
Qy	2641	aaatcaatgtcaagtaccaaaaatgttaatatgtatgttcaatttaactctgctgagactt	2700
Db	2641	AAAATCAATGTCAAGTACCAAAATGTTAATGTATGTGTCTCATTTAACTCTGCTGAGACTT	2700
Qy	2701	tcagtgcaactgtatataagaagctctaaaaacacacactaaagaaaaagaatcgaattttcag	2760
Db	2701	TCAGTGCACTGTATATAGAAGTCTAAACACACACTAAGAGAAAAAGATCGAATTTTTCAG	2760
Qy	2761	atgattcggaattttcattccaggtatttgaatagtagacatatatatgtatatatacat	2820
Db	2761	ATGATTTCGGAAATTTTCATTCAGGTATTTGTATAATAGTACATATATATGTATATACAT	2820
Qy	2821	caactcctattctcttaattttttaaagttaaaactgtaactggcagtaagctttttttgatca	2880
Db	2821	CACCTCCTATTCTCTTAATTTTGTAAATGTTAACTGGCAGTAAGTCTTTTGTGATCA	2880
Qy	2881	ttcccttttccatataggaagaaataaattttgaagtgggcagatgagtttatcatgtcagt	2940
Db	2881	TTCCCTTTTCCCATATAGGAACATAAATTTGAAGTGGCCAGATGAGTTTATCATGTGAGT	2940
Qy	2941	gaaaaataattaccccaaaatgccacagtaacttaacgattcttcaactctctggggttt	3000
Db	2941	GAAAAATAATTACCCAAATGCCACCAGTAACTTAACGATTTCTTCACTTCTTTGGGGTTT	3000
Qy	3001	tcagtgatgaacctaaactcccccaacacatctccctcccacattgtcacccatttcaaaag	3060
Db	3001	TCAGTATGAACCTTAACCTCCCAACCAACATCTCCCTCCACATTTGTACCACTTTTCAAG	3060
Qy	3061	ggcccaacagtgacttttgcctgggcattttcccagatgttttaacagactgtgagtaacagcag	3120
Db	3061	GGCCCAACAGTACTTTTGTCTGGGCATTTTCCCAGATGTTTACAGACTGTGAGTACAGCAG	3120
Qy	3121	aaaatttttaactagtggtgtgtatataataaacaattgtgtaataatttcttttagccc	3180
Db	3121	AAAATCTTTTACTAGTGTGTGTGTATATATATAAACAAATTTGTAAATTTCTTTTAGCCC	3180
Qy	3181	attttctagactgtctctggtggaataattttgtgtgtgtatgtatgcattgtgtgtgat	3240
Db	3181	ATTTTCTAGACTGTCTCTGTGGAATATATTCTGTGTGTGTATATATGATGATGTGTGTGAT	3240
Qy	3241	ggtatgtatgattttaatcctaataaattgtgcccgcagttgtgccaagtgacata	3300
Db	3241	GGTATGTATGGATTTAATCTAATTAATTTGTGCCCGCAGTTGTGCCAAAAGTGCAATA	3300
Qy	3301	gtctgagctaaaaatcctagtgattgttcatcatgacaacctcctcagtcacatttlaacc	3360
Db	3301	GTCTGAGCTAAAAATCTAGGTGATTTGTTCATCATGACAACCTGCTCAGTCCATTTTAACC	3360
Qy	3361	tgtagcaacctctctgcattcataaattcttgttaactgtttaccattacaaatgggata	3420
Db	3361	TGTAGCAACCTTCTGCATTCATAAATCTTGTAAATCATGTGTACCAATACAAATGGGATATA	3420
Qy	3421	agaggcagcgtgaaacagacagatgagctgtggactagcaataaggggttttctgtgtggt	3480
Db	3421	AGAGGCAGCGTGAACAGACATGAGCTGTGGACTAGCAATATAGGGGTTTTTGTGGTGTGT	3480
Qy	3481	tgggttgataaagcagctatttggggtcatattgtttccctgtgctggagcaaaagtcattca	3540
Db	3481	TGGTTTCAATAAGCAGTATTTGGGGTCAATATTGTTTCCCTGTGCTGGAGCAAAAGTCATTA	3540
Qy	3541	cactttgaagatttatattgttcttatccctcaattcaatgtgtgagtaaatggccaggt	3600
Db	3541	CACCTTTGAAGTATTATATTGTTCTTATCTCAATTCATGTGTGTGATGAATTTGCCAGGT	3600
Qy	3601	tgtctgatatttcttccagacttcgccacagagattcgtgataataaaataggtgaagaca	3660
Db	3601	TGTCTGATATTTCTTTTCAGACTTCGCCACAGATTTGCTGATATAAATAATTAGTAAGATA	3660
Qy	3661	atttctggggccatatatttagacacaggttaaaataacacaggttccagttcgttggaattg	3720
Db	3661	ATTTGTGGGCCATATTTTAGGACAGGTAAAAATAACATCAGGTTTCCAGTTGCTTGAATTG	3720

```

QY 3721 caaggctaagaagtactgcccttttggtagcagtcacaaatctatttccactggcg 3780
Db 3721 CAAGGCTAAGAAGTACTGCCCTTTGGTAGCAGTCAAAATCTATTATCCACTGGCG 3780
QY 3781 catcatgtagtgatataatgcctataataaagcagtaggttccacacattttgttag 3840
Db 3781 CATCATGTCAGTATATGCTTATATAGCCTATAATAGCCATAGGTTTCACACCAATTTGTTAG 3840
QY 3841 acaattgtcttttttcaagatgctttgttttttttcatataaataaaatgcattttataa 3900
Db 3841 ACAATTGCTTTTTCAGATGCTTTGTTCTTTCATATGAAAAAATGCAATTTATATAA 3900
QY 3901 attcagaagtcagtagatttctgaaggcgcgaacgtgcatattttatttgtagactggtaag 3960
Db 3901 ATTGAGAAAGTCATAGATTTCTGAAGCGCTCAACGTGCATTTTATTATGGACTGGTAAG 3960
QY 3961 taactgtgtttactagcaggaataattcccaatttctaccctttactatacttttcaaca 4020
Db 3961 TAACGTGTGTTTACTAGCAGGAATATTTCCAATTTCTACTTACTATCTTTTCAACA 4020
QY 4021 agtaacttttagaataagccagaagccaagggcctgagttggcagtgggccataaagt 4080
Db 4021 AGTAACTTTGTAGAAATGAGCCAGAGCCAGGCCCTGAGTTGGCAGTGGCCCAATAGTG 4080
QY 4081 taaataaaagtttacagaaacatt 4105
Db 4081 TAAATAAAAGTTTACAGAAACCTT 4105

RESULT 3
LOCUS HSET1R 4105 bp mRNA linear PRI 23-MAR-1993
DEFINITION H.sapiens mRNA for endothelin-1 receptor.
ACCESSION X61950
VERSION X61950.1 GI:288312
KEYWORDS endothelin-1 receptor.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 4105)
AUTHORS Hosoda,K., Nakao,K., Hiroshi-Arai, Suga,S., Ogawa,Y., Mukoyama,M.,
Shirakami,G., Saito,Y., Nakanishi,S. and Imura,H.
TITLE Cloning and expression of human endothelin-1 receptor cDNA
JOURNAL FEBS Lett. 287 (1-2), 23-26 (1991)
MEDLINE 91348221
FEATURES
    source          Location/Qualifiers
    1..4105
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    485..1768
    /codon_start=1
    /product="endothelin-1 receptor"
    /protein_id="CAA43953.1"
    /db_xref="GI:288313"
    /db_xref="SWISS-PROT:P25101"
    /translation="METLCRASPMALVGVCSIDNPERYSTNLNHDVDPETPRGTE
    LSPLETHQTNLVPNSGMHNYCPQTKITSAPKYNIVISCTIFVGMGNATLL
    RIYQKMRNGPNALIASALGLIYVIDLPIVFNKLLAGRPWFHDNDFVFLKL
    FPLQKSSVGTVLNICALSVRYAVASWSRVQIGIPLVTAIEIYSIWILSLAI
    PEAFGVNVPFVYRGQKTCMLNATSKFMFYODVKDWLFGFYFCMPLVCTAIFT
    LMTCEMLNRRGSLRIALSEHLKORREAVTVCLVVFALCWPLHLRLILKKTYN
    EMDKNRCELLSFLLLMDYIGINLATMNSCINPIALYFVSKFKNCFOSCLCCCCYQSK
    SLMTSPVMNTGTSQWKNDQNNHNDKSHKDSMN"
BASE COUNT 1138 a 859 c 845 g 1263 t
ORIGIN
Query Match 100.0%; Score 4105; DB 9; Length 4105;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4105; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 gaattcggcgccctcttgcgtcccgagatggagtggaaggtctgagcgttgggagg 60

```

```

Db 1 GAATTGCGCGCGCCTCTTTCGGGTCCAGAGTGGAGTGAAGGCTCGAGCTTTGGGAGG 60
QY 61 agacggggagagacagactgagagcggtgttctctccgagagtttttttttttctgagagcct 120
Db 61 AGACGGGGAGGAGACAGACTGGAGCGGTGTTCTCCCGAGTGTTCCTTTTTCGTCGAGCCCT 120
QY 121 cgcgcgcgcgtacagtcacccgcgcgtgctgacgatttggagagcggtggagagcctt 180
Db 121 CCGCGCGCGGTACAGTCATCCGCTGCTGACGATTGTGGAGAGGGGTGGAGAGGCTT 180
QY 181 catccatccaccccggtcgtccgcgggattggggtcccgagcagacacccctcccgaggaa 240
Db 181 CATCATCCACCCCGTCTGCGCGGGATTGSGGTCCCGAGCAGACCTCCCCCGGAGAA 240
QY 241 gcagtcgccaggaagttttctgaagccggggaagctgtgcagccgaagcgcgcgcgcgc 300
Db 241 GCAGTGCACGAGAAAGTTTCTGAAGCCGGGGAAGCTGTGCAGCCGAGCGCGCGCGCGC 300
QY 301 cggagccggggacacccgcgcacccctccgcgcacccacccctcgtcttctcggcttcctc 360
Db 301 CGGAGCCCGGGACACCGGCCACCTCCGGCCACCCACCCCTCGCTTCTCCGGCTTCCTC 360
QY 361 tggccagcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 420
Db 361 TGGCCAGCGCGCGCGCGGACCCGCGGAGCTGTGCGCGCAGCTCCACCGAGTCCACGGTGA 420
QY 421 aaaaagtgaagtgtaaaagcagcagaagtgcaataagagatatcttctccaaattgacct 480
Db 421 AAAAAAGTGAAGGTGTAAGAACGACGACAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 480
QY 481 caagatgaacccctttgcctcagggcacccttttgcctgagcagcagcagcagcagcagc 540
Db 481 CAAGATGAAGAACCCCTTTGCCTCAGGCACTCCCTTTGGTGGCACTGGTGGATGTGAAT 540
QY 541 cagtgataatcctgagagatagacagcacaatctaaagcaatcagtgagagatttccaccac 600
Db 541 CAGTGATAATCCTGAGAGATACAGCACAAATCTAAGCAATCATGTGATGATGATGATGATG 600
QY 601 ttttcgtggcagagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 660
Db 601 TTTTCGTGGCAGAGCTCAGCTTCCTGTTTACCTACCTACCTACCTACCTACCTACCTACCT 660
QY 661 acccagcaatgctcaatgacacacacacacacacacacacacacacacacacacacacac 720
Db 661 ACCCAGCAATGGCTCAATGACACAACTATTGCCCACAGCAGACTPAAATTTACTTCACTTT 720
QY 721 caaatacattacacactgtgatatcttctactattttctcgtgggaatggtggggaatgc 780
Db 721 CAAATACATTAACTGTGATATCTGTACTATTTCATCGTGGGAATGGTGGGAATGC 780
QY 781 aactctgctcaggtacattaccagaacaaatgtaggaatggtggcgaacgcgtgat 840
Db 781 AACTCTGCTCAGGATCATTTTACCAGAAATGATGAGGAATGGCCCCAACACGCGTGAT 840
QY 841 agccagcttgccttgcagacattatctatggtgcttgcctcctcctcctcctcctcctcct 900
Db 841 AGCCAGCTTGCCTTGGAGACCTTATCTATGTGGTGCATGTGATCTCCCTATCAATGATT 900
QY 901 taagctgctggtggcgctggccttttgcacacacacacacacacacacacacacacacac 960
Db 901 TAAGCTGCTGGCTGGCGCTGGCCCTTTTGATCACAATGACTTTGGCGTATTTCCTTTTGCA 960
QY 961 gctgttccctcttttgcagaagtcctcgtggggatcacccgtcctcctcctcctcctcctc 1020
Db 961 GCTGTCCCTCTTTTGCAGAAAGTCTCGGTGGGATCACCGTCTCAACCTCTCGCGCTCT 1020
QY 1021 tagtttacaggttacagagcaggttgcctcctcgtgagtcgtgttcaggggaattgggattcc 1080
Db 1021 TAGTGTTCACAGGTACAGAGCAGTGTCCCTCCCTGAGTCTGTTTCAGGGAATTTGGGATTC 1080
QY 1081 ttgtgtaactgacattgaaattgctccatctggacccctgctcttcttcttctcctgacattcc 1140

```

Db	1081	TTTGGTAACCTGCCATTCAAATTTGCTCCATCTGGATCCTGCTCCTTTATCTCTGGCCATTCC	1140
QY	1141	tgaagcagattggttcgtcatgtaaaccttttgaataataggggtgaacagacataaaacctg	1200
Db	1141	TGAAGCGATTGGGTTGCTCATGCTACCTTTGAAATAGGGTCAACAGCATAAACCTG	1200
QY	1201	tatgcacaaatgccacatcaaaattcatggagttctaccagaatgtaaaaggactgggtgct	1260
Db	1201	TATGCTCAATGGCCACATCAAAATTCATGGAGTTCTACCAAGATGTAAAGGACTGGTGGCT	1260
QY	1261	cttcgggttctatttctgtatgccccttgggtgctgactggcatctctacacccctcatgac	1320
Db	1261	CTTCGGGTCTATTCTGTATGCCCTTGGTGTGCATGCGCATCTTCTACACCCCTCATGAC	1320
QY	1321	ttgtgagatgtgaacagaaggaaatggcagcttgagaattggccctcagtgaaacatcttaa	1380
Db	1321	TTGTGAGATGTTGAACAGAAGGAATGCGACGTTGAGAAATGGCCCTCAGTGAACATCTTAA	1380
QY	1381	gcagcgtcgaagaagtggcaaaaacagtttctcgttgggttggtaatttttgcctcttgcgtg	1440
Db	1381	GCAGCGTCGAGAAGTGGCAAAAACAGTTTCTCCTTGGTGTAAATTTTGCCTCTCTTGC	1440
QY	1441	gttccctcttcacttaagccgtatattgaagaaaactgtgtataacgaaatgacaagaa	1500
Db	1441	GTTCCCTCTTCACCTTAAGCCGTATATTGAAGAAACTGTGTATAACGAATGGACAAGAA	1500
QY	1501	ccgatgtgaattacttagtttcttactgctcatggattacatcgggtattaaacttggcaac	1560
Db	1501	CCGATGTGAATTACTTAGTTTCTTACTGCTCATGGATTACATCGGTATTAACTTGGCAAC	1560
QY	1561	catgaattcatgtataaaacccctagctctgtattttgtgagcaagaaattttaaattg	1620
Db	1561	CATGAATTCTATGTATAAACCCCTAGCTCTGTATTGTGTGAGCAAGAAATTTAAAAATTG	1620
QY	1621	ttccagtcgctgctgctgctgtgtaccagtcacaaagtctgatgaacctcgtgctccc	1680
Db	1621	TTTCCAGTCATGCTCTGCTGTGCTGTACCAAGTCCAAAGTCTGATGACCTCGGTCCC	1680
QY	1681	catgaacgaacaaagcatccagtggaagaaacacacgatacaaaacacacacacagacg	1740
Db	1681	CATGAACGAACAAAGCATCCAGTGGAGAACACACGATCAAAACACCAACACACAGACCG	1740
QY	1741	gaacagccataaagacagcatgaactgaccaccccttagaagaactcctcgtgtaacctcat	1800
Db	1741	GACAGCCATAAAGGACAGATGAACATGACCACCTTAGAAGCACTCCTCGGTACTCCCAT	1800
QY	1801	aatcctctcggagaaaaaatcaaaaggcaactgtgactccgggaatctctctctgctgc	1860
Db	1801	AATCCTCTCGGAGAAAAAATCAAGGCAACTGTGACTCCGGGAATCTCTCTCTGATC	1860
QY	1861	cttcttctcttaattcactccacaccccaagaagaatgtcttccaaaacccgcaagtaga	1920
Db	1861	CTTCTTCTCTTAATCTACTCCACACCAAGAAGAAATGCTTTCCAAAACCGCAAGGTAGA	1920
QY	1921	ctggtttatccaccccaacacatacgaatcgtactcttctttaaattgatctaaattacata	1980
Db	1921	CTGGTTTATCCACCCCAACATCTACGAATCGTACTCTTTAATTGATCTAATTTACATA	1980
QY	1981	tctcgtgctgtatttcagcactaaaaatggtggagctgggggagaaatgaagactgtt	2040
Db	1981	TTCCTGCTGTTGTATTACACACTTAAAAATGTTGGGAGCTGGGGGAGAAATGAAGCTGTT	2040
QY	2041	aaatgaacccagaaggatatattactacttttgcatagaataagagcttttcaagtatagg	2100
Db	2041	AAATGAACCCAGAAGGATATTACTACTTTTGCATGAAATAGAGCTTTCAAGTACATGG	2100
QY	2101	ctagcttttatggcagctctcgtgaatgttcaatgggaactgggtcaccatgaaactttg	2160
Db	2101	CTAGCTTTTATGGCAGTCTCTGTGAATGTTCAATGGGAACCTGGTACCATGAAACTTTAG	2160
QY	2161	agattaacgacagaagatttctactttttttaaagtgatttttttgccttcagcccaaacaca	2220
Db	2161	AGATTAAACGACAAGATTTTCTACTTTTAAAGTGAATTTTGTCTCTTCTCAGCCAAACACA	2220
QY	2221	atatggctcaggtcactttttttaaagtgtcatttgggtggtccagtagttttttaactgca	2280
Db	2221	ATATGGGCTCAGGTCACTTTTATTGAAATGTCATTTGGTGCCAGTATTTTAACTGCA	2280
QY	2281	taatagcctaaactgatatttgaacttattacacatagtttggaaataaaaaaagacaaa	2340
Db	2281	TAATAGCCTAACATGATATTGTGAACCTATTACACATAGTTTGAATAAAAAAAGACAAA	2340
QY	2341	aataglattcaggtgagcaattagattagtttttccacgtcactatttttttttaaa	2400
Db	2341	AATAGTATTACGCTGAGCAATTAGATTAGTATTTTCACCTCACTATTATTATTTTAA	2400
QY	2401	acacaaattctaaagctacacaaaactactacagggcccttaaaagcagctgtgatgacaca	2460
Db	2401	ACACAAATCTAAAGCTACAAATACTACAGGCCCTTAAGACACAGCTGTATATTTTAA	2460
QY	2461	tttggcagtttaatagatgttactcaaaagaattttttaagaactgtatttttttttaa	2520
Db	2461	TTTGGCAGTTTAATAGATGTACTCAAAAGAATTTTAAAGAACTGTATTTTATTTTAA	2520
QY	2521	atggttttttattacaaggaccttgaacatgtttgtatgttaaaattcaaaagttaatgc	2580
Db	2521	ATGGTGTTTTATTACAGGGGACCTTGAACATGTTTGTATGTATAAATTCAAAAAGTAATGC	2580
QY	2581	ttcaatcagatagttctcttttcaaaagtccaactactgttttcatgtataaattttgtatga	2640
Db	2581	TTCAATCAGATAGTCTTTTTCACAGTTCATACTGTTTTCATGTAAATTTTGTATGA	2640
QY	2641	aaatcaatgtcaagtaccacaaatgttaattgtatgtgtcttaactctgctgagactt	2700
Db	2641	AAAAATCAATGTCAAGTACCAAAATGTTAATGTATGTCTAATTTAACTCTGCCTGAGACTT	2700
QY	2701	tcagtcactgtatatagaagctcaaaacacacacacacacacacacacacacacacacac	2760
Db	2701	TCAGTGCACCTGTATAGAAGTCTAAACACACCTAAGAGAAAGAAAGTCAATTTTCAG	2760
QY	2761	atgattcggaaatttttcattcaggtatttgaatagtgacataataatgtataacatat	2820
Db	2761	ATGATTCGGAATTTTTCATTTCAGGTATTGTAATAGTGCATATATATGATATACATAT	2820
QY	2821	caactcctattctcttaatttttggtaaaatgttaactggcagtaagcttttttggatca	2880
Db	2821	CACCTCTATCTCTCTTAATTTTGTAAAAATGTTAACTGGCAGTAGTCTTTTTCATCA	2880
QY	2881	ttcccttttccatcataggaacataatttgaagtgccagatgaagttatcatgacgt	2940
Db	2881	TTCCCTTTTCCATATAGGAACATAATTTTGAAGTGGCCAGATGAGTTTATCATGTGAGT	2940
QY	2941	gaaaaataatttcccccaaatgccaccagtaacttaacgattcttcactctcttggggtt	3000
Db	2941	GAAAAATAATTACCCCAAAATGCCACCAGTAACCTTAACGATTCTTCACTTCTTGGGGTTT	3000
QY	3001	tcagtagaaccttaactcccccccccaacacatctccctccacattgtcaccatttcaag	3060
Db	3001	TCAGTATGAACCTTAACCTCCCAACATCTCCCTCCACATTTGTCACCATTTCAAAG	3060
QY	3061	ggccacagtgacttttgcctgggcatcttcccagatgtttacagactgtgagtacagcag	3120
Db	3061	GGCCCAAGTGACTTTTGTCTGGGCATTTTCCCAAGATGTTTACAGACTGTGAGTACAGCAG	3120
QY	3121	aaaaatttttactagtggtgtgtgtatatataaacaattgttaaaattcttctttagccc	3180
Db	3121	AAAAATCTTTTACTAGTGTGTGTGTATATATAAACAATGCTAAATTTCTTTTAAAGCCC	3180
QY	3181	attttctagactgtctgtggaatataatttgtgtgtgtgtatgatcatgtgtgtgat	3240
Db	3181	ATTTTCTAGACTGTCTGTGGAATATATTGTGTGTGTGTATATGATGTGTGTGAT	3240
QY	3241	ggtatgtatggatttaattcatttaataatgtgccccgcaggtgtgccccaaagtgcata	3300
Db	3241	GGTATGTATGGATTTTAAATCTAATTAATTTGTGCCCGCAGTTGTGCCAAAGTGCATA	3300

QY	3301	gtctagctaaatctaggtgattgttcacatgatgacaaacctgacctgaattttaaacc	3360
Db	3301	gtctagctaaatctaggtgattgttcacatgatgacaaacctgacctgaattttaaacc	3360
QY	3361	tgtagcaacctcttcattcataatcttgtaatactgttaccattacaaatgggatata	3420
Db	3361	tgtagcaacctcttcattcataatcttgtaatactgttaccattacaaatgggatata	3420
QY	3421	agaggcagctgtaagcagatgagctgtggactgacaaataggggttttgggttggt	3480
Db	3421	agaggcagctgtaagcagatgagctgtggactgacaaataggggttttgggttggt	3480
QY	3481	tggtttgataaagcagatattgggtgcatattgttctctgtctgagcaaaagtcatt	3540
Db	3481	tggtttgataaagcagatattgggtgcatattgttctctgtctgagcaaaagtcatt	3540
QY	3541	cactttgaagtattatattgtcttctcctcaatcaatgctgtaataaaatggcaggt	3600
Db	3541	cactttgaagtattatattgtcttctcctcaatcaatgctgtaataaaatggcaggt	3600
QY	3601	tgctgataattcttctcagacttcgacagacagattgctgataataaattaggttaagata	3660
Db	3601	tgctgataattcttctcagacttcgacagacagattgctgataataaattaggttaagata	3660
QY	3661	attgttggtgacatttttagacagaggttaaaataacatcaggttccagttgctgaattg	3720
Db	3661	attgttggtgacatttttagacagaggttaaaataacatcaggttccagttgctgaattg	3720
QY	3721	caaggctaagagtaactgctcttctgtgttagcagtcacaaatctatttccactggcg	3780
Db	3721	caaggctaagagtaactgctcttctgtgttagcagtcacaaatctatttccactggcg	3780
QY	3781	catcatatgcagtgatataatgcctataataagccataggttcacacatttttttag	3840
Db	3781	catcatatgcagtgatataatgcctataataagccataggttcacacatttttttag	3840
QY	3841	acaattgtcttttttcaagatgcttcttcttcttcaatgaataaaatgcattttataa	3900
Db	3841	acaattgtcttttttcaagatgcttcttcttcttcaatgaataaaatgcattttataa	3900
QY	3901	attcagaagtcagatattctgaagcgctcaactgcatattttatgtgactgtaag	3960
Db	3901	attcagaagtcagatattctgaagcgctcaactgcatattttatgtgactgtaag	3960
QY	3961	taactgtgttactagcaggaataatttccatttctacatttcttactacatttcaaca	4020
Db	3961	taactgtgttactagcaggaataatttccatttctacatttcttactacatttcaaca	4020
QY	4021	agtaactttgtagaatagcaggaagccagggccctgagttgagtgcccaataagtg	4080
Db	4021	agtaactttgtagaatagcaggaagccagggccctgagttgagtgcccaataagtg	4080
QY	4081	taaaataaagtttacagaaacctt	4105
Db	4081	taaaataaagtttacagaaacctt	4105
RESULT	4		
LOCUS	S57498	4105 bp	linear
DEFINITION	endothelin ET-A receptor [human, mRNA, 4105 nt].		
ACCESSION	S57498		
VERSION	S57498.1	GI:298319	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4105) Arai, H., Nakao, K., Hosoda, K., Ogawa, Y., Nakagawa, O., Komatsu, Y. and Imura, H.		
TITLE	Molecular cloning of human endothelin receptors and their expression in vascular endothelial cells and smooth muscle cells		

Jpn. Circ. J. 56, 1303-1307 (1992)	
93180293	
1291713	
GenBank staff at the National Library of Medicine created this entry [NCBI gibbsq 128422] from the original journal article.	
This sequence comes from Fig. 5.	
Location/Qualifiers	
1..4105	
/organism="Homo sapiens"	
/db_xref="taxon:9606"	
485..1768	
/note="This sequence comes from Fig. 5; conceptual translation presented here differs from translation in publication; ET-AR"	
/codon_start=1	
/product="endothelin ET-A receptor"	
/protein_id="AAB25530.2"	
/db_xref="GI:18390352"	
/translation="METICLRASFWLALVGVISDNPERYSTNLSNHVDFTFRGTE LSLVTHPTNLVPSNGSHNYCPQOTKITSAFYINTVICTIFIVGMVGNATLL RIYONKMRNPNALIASLADLIYVVDIPNVFKLAGRWPFHDHDFGVFLCKL FPEIQKSSGIVTLNLCALSDVRYRAVSWRVQIGIPLVTAIEIVSWILSFLILAI PEAGFVMVPEYGEQHKTMATSKMEFYQDKDWMLEFGFVCMPLVCTAIFYT LMTCEMLNRXNGSLRIALSEHLKQREYAKTVFLVIFALCWFFLHLKRIKTVYN EMDKNRCLESLFLLDYIGINLATMNSCINPIALYFVSKFKNCFQSCLCGCCQYQSK SLMTSPVPMNGTSIQWNHNDQNNHNDRSSHKDSMN"	
BASE COUNT	1138 a 859 c 843 g 1263 t 2 others
ORIGIN	
Query Match	100.0%; Score 4103; DB 9; Length 4105;
Best Local Similarity	100.0%; Pred. No. 0;
Matches 4103; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	1 gaattcgccgagcgcctcttgcgggtccagagtggaagtgagagctcggagcttggagg 60
Db	1 GAATTCGGCGGCGCCTCTTTCGGGTCCAGAGTGGAGTGGAGCTCTGGAGCTTGGAGG 60
QY	61 agacgggagggacagactgagggcgtgttcctccgagtttcttcttcgtgcgagccct 120
Db	61 AGACGGGGAGGACACACTGGAGCGGTGTCTCCCGAGATTCTTTCTGTCGGAGCCCT 120
QY	121 cgcgcgcgcgtacagtcacccgcctcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 180
Db	121 CGCGCGCGGTACAGTCATCCGCTGGTCTGACGATTGTGGAGAGCGGTGGAGAGCTT 180
QY	181 catccatccacccgc 240
Db	181 CATCCATCCACCCCGGTCGTCGCGGGGATTGGGGTCCCGAGCACACCTCCCGCGGAGAA 240
QY	241 gcagtgcccgaggaagtcttctgaagcgggggaagctgtgcagccgaagcgcgcgcgcgc 300
Db	241 GCAGTGCCCGAGGAGTTTCTGAGCCGGGGAAGCTGTGACCCGAAGCGCGCGCGCGC 300
QY	301 cggagcccggggacacccgc 360
Db	301 CGGAGCCCGGGACACCGGCCACCTCCGCGCGCACCCACCCACCTCCCGCGCTTCCCTC 360
QY	361 tggcccgagc 420
Db	361 TGGCCGAGCGCGCGCGGACCCGCGAGCTGTCTGCGCAGCCGAGCTCCACCGTGAATA 420
QY	421 aaaaagtgaagtgtaaaagcagcaagtcgaataaagagatatatttcccaaatggcct 480
Db	421 AAAAAGTGAAGTGTAAAGAGCAGCAAGTGCACAAAGAGATATTTCTCAAAATTTGCCT 480
QY	481 caagatggaaccccttgcctcagggcctccttttggctggcactggctggatgtgtaata 540
Db	481 CAAGATGGAACCCCTTTCCTCAGGCGCATCTTTTGGCTGGCACTGTTGGATGTGTAAT 540
QY	541 cagtgataatcctgagagatacagcacaaatcctaagcaatcattgtgagattaccacac 600
Db	541 CAGTGATAATCCTGAGAGATACAGCACAAATCTAAGCAATCATGTGGATGATTTCACCAC 600

QY 601 ttttcgtggcagagctcagcttctcgttggttaacactcatcaaaacccactaatttggctcc 660
DB 601 ttttctggggcacagagctcagcttctcgttggttaacactcatcaaaacccactaatttggctcc 660
QY 661 acccagcaatggctcgaatcacactattgcccacagcagactaaaattacttcagcttt 720
DB 661 acccagcaatggctcgaatcacactattgcccacagcagactaaaattacttcagcttt 720
QY 721 caaatacattaaactgtgatatcttctgtactatttttctcgtgggaatggtgggaatgc 780
DB 721 caaatacattaaactgtgatatcttctgtactatttttctcgtgggaatggtgggaatgc 780
QY 781 aactctgctcagatcattttacagaaacaaatgatgaggaatggtggcccaacgcgtgat 840
DB 781 aactctgctcagatcattttacagaaacaaatgatgaggaatggtggcccaacgcgtgat 840
QY 841 agccagcttggccttggagaccttatctatggtggtcattgatctccctacatgaattatt 900
DB 841 agccagcttggccttggagaccttatctatggtggtcattgatctccctacatgaattatt 900
QY 901 taagctgctggctggcgtggccttttgcacaaatgacttttggcgattttcttggcaa 960
DB 901 taagctgctggcgtggcgtggccttttgcacaaatgacttttggcgattttcttggcaa 960
QY 961 gctgttcccttttttgcagaagctcctggttggggatacccgctcctcaacccctcgcctct 1020
DB 961 gctgttcccttttttgcagaagctcctggttggggatacccgctcctcaacccctcgcctct 1020
QY 1021 tagtctgacagctacagagcagcttgcctcctgagctggttccaggaatggtgattcc 1080
DB 1021 tagtctgacagctacagagcagcttgcctcctgagctggttccaggaatggtgattcc 1080
QY 1081 tttggtaactggcattgaaattgtctccatctggtcctgcttcttctcctggtccattcc 1140
DB 1081 tttggtaactggcattgaaattgtctccatctggtcctgcttcttctcctggtccattcc 1140
QY 1141 tgaagcgattggcttgcctgcatggtacccttttgaatataggggtgacagcagataaacctg 1200
DB 1141 tgaagcgattggcttgcctgcatggtacccttttgaatataggggtgacagcagataaacctg 1200
QY 1201 tatgtccaatggccacataaaattcagtgagttctcaccagatgtaaaaggactggct 1260
DB 1201 tatgtccaatggccacataaaattcagtgagttctcaccagatgtaaaaggactggct 1260
QY 1261 cttcgggttctatttctgtatgccccttggctgctgctgctgcttctcaccctcctgac 1320
DB 1261 cttcgggttctatttctgtatgccccttggctgctgctgctgcttctcaccctcctgac 1320
QY 1321 tctgagagatttgaacagagaaggaatggcagcttgagaattgcccctcagtgaaactcttaa 1380
DB 1321 tctgagagatttgaacagagaaggaatggcagcttgagaattgcccctcagtgaaactcttaa 1380
QY 1381 gcaagctcagaagtgccaaaacagtttctgcttggcttgggttgaatttctgcttggctg 1440
DB 1381 gcaagctcagaagtgccaaaacagtttctgcttggcttgggttgaatttctgcttggctg 1440
QY 1441 gttccctcttccacttaagcgttatattgaagaaaactgtgtataacgaaatggacaagaa 1500
DB 1441 gttccctcttccacttaagcgttatattgaagaaaactgtgtataacgaaatggacaagaa 1500
QY 1501 ccgagtggaattacttagtttcttactgctcattgattacacgcggtattaaacttggcaac 1560
DB 1501 ccgagtggaattacttagtttcttactgctcattgattacacgcggtattaaacttggcaac 1560
QY 1561 catgaattcatgtataaaccccatagctctgatttttggagcaagaaatttaaaaattg 1620
DB 1561 catgaattcatgtataaaccccatagctctgatttttggagcaagaaatttaaaaattg 1620
QY 1621 tttccagctcagctcctgctgctgtgtaccagttccaaagctcgtgacacctcggtccc 1680
DB 1621 tttccagctcagctcctgctgctgtgtaccagttccaaagctcgtgacacctcggtccc 1680

QY 1681 catgaacggaacagcatccagtggaagaaacccagctcaaaaacacacacacagaccg 1740
DB 1681 catgaacggaacagcatccagtggaagaaacccagctcaaaaacacacacacagaccg 1740
QY 1741 gacagccataaagacagcatgaactgaccaccccttgaagcactcctcgttactcccat 1800
DB 1741 gacagccataaagacagcatgaactgaccaccccttgaagcactcctcgttactcccat 1800
QY 1801 aatcctctcggagaaaaaatacacaaggaactgactcgggaatctcttctctgac 1860
DB 1801 aatcctctcggagaaaaaatacacaaggaactgactcgggaatctcttctctgac 1860
QY 1861 cttcttcttaatttactcctccacacaccccaagaataatgcttccaaaacgcgaagtaga 1920
DB 1861 cttcttcttaatttactcctccacacaccccaagaataatgcttccaaaacgcgaagtaga 1920
QY 1921 ctggtttatcccccacacacatctacgaatcgtacttcttcttaattgacttaattacata 1980
DB 1921 ctggtttatcccccacacacatctacgaatcgtacttcttcttaattgacttaattacata 1980
QY 1981 tctcgcgtgtgtattcagcactaaaaatggtggagctggggagaaatgaagactgtt 2040
DB 1981 tctcgcgtgtgtattcagcactaaaaatggtggagctggggagaaatgaagactgtt 2040
QY 2041 aatgaacacagagaagatatattactacttcttgcataaaaaatagagcttcaagtacatgg 2100
DB 2041 aatgaacacagagaagatatattactacttcttgcataaaaaatagagcttcaagtacatgg 2100
QY 2101 ctggttttatggcagcttctggtgaatgttcaatgggaactggtcaccatgaaactttag 2160
DB 2101 ctggttttatggcagcttctggtgaatgttcaatgggaactggtcaccatgaaactttag 2160
QY 2161 agatttaacacagaagatttctacttctttaaagtgatttttctcctcagcacaacaca 2220
DB 2161 agatttaacacagaagatttctacttctttaaagtgatttttctcctcagcacaacaca 2220
QY 2221 atagggctcaggtcacttcttattgaaatgctatttgggcagtagtatttttaactgca 2280
DB 2221 atagggctcaggtcacttcttattgaaatgctatttgggcagtagtatttttaactgca 2280
QY 2281 taataccttaacatgattatttgaacttatttcaacatagtttggaaaaaaagacaaa 2340
DB 2281 taataccttaacatgattatttgaacttatttcaacatagtttggaaaaaaagacaaa 2340
QY 2341 aatagttacagtgagcaattagatttatttccacgctcactatttattttttaa 2400
DB 2341 aatagttacagtgagcaattagatttatttccacgctcactatttattttttaa 2400
QY 2401 acacaaattcctaaagctacacaaaatactacagggccctttaaagcagctctgtagacaca 2460
DB 2401 acacaaattcctaaagctacacaaaatactacagggccctttaaagcagctctgtagacaca 2460
QY 2461 tttggcagtttaataagatgttactcaaaataatttttaagaactgatttttttttaa 2520
DB 2461 tttggcagtttaataagatgttactcaaaataatttttaagaactgatttttttttaa 2520
QY 2521 atgggtgttttatacagaagccttgaacatgttttggatgttttaaatcaaaagtagtc 2580
DB 2521 atgggtgttttatacagaagccttgaacatgttttggatgttttaaatcaaaagtagtc 2580
QY 2581 tccaatcagatagtttcttttccaaagtccaatactggttttctcatgtaaaatttctgtaga 2640
DB 2581 tccaatcagatagtttcttttccaaagtccaatactggttttctcatgtaaaatttctgtaga 2640
QY 2641 aaaaataatgctcaagtacccaaaatgttaatgtatgtgtcattttaaactcgtcgtgagactt 2700
DB 2641 aaaaataatgctcaagtacccaaaatgttaatgtatgtgtcattttaaactcgtcgtgagactt 2700
QY 2701 tcagtgactgtatagagcttaaaacacacacccctaaagagaaaaagatcgaaattttcag 2760
DB 2701 tcagtgactgtatagagcttaaaacacacacccctaaagagaaaaagatcgaaattttcag 2760
QY 2761 atgattcggaaatttttcttccagggtatttggtaatagtagacatatatatgatacatat 2820


```

Db 2761 ATGATTCGGAAATTTTCATTCAGGTATTGTAAATAGTGACATATATATGATATACATAT 2820
Qy 2821 caccctcctattctcttaattttttaaataatttaaacgtgaagcagtaagcttttttttgatca 2880
Db 2821 CACCTCCCTATTCTCTTAATTTTGTAAATGTTAACTGGCAGTAAGCTCTTTTGTGATCA 2880
Qy 2881 ttcccttttccatagaaacataattttgaagtggccagatgagtttatcatgtcagt 2940
Db 2881 TTCCCTTTTCCATATAGGAACATAATTTTGAAGTGGCCAGATGAGTTATCATGTCAGT 2940
Qy 2941 gaaaaataattaccacaaaatgcaccsgtaacttaacgattcttcacttcttggggttt 3000
Db 2941 GAAAAATAAATTAACCCACAAATGCCACCAGTAACCTTAAGGATTTCTTCACTTCTTGGGGTT 3000
Qy 3001 tcagtatgaacctaacctcccccacacatctccctcccacattgtcaccaatttcaag 3060
Db 3001 TCAGTATGAACCTAACTCCCCACCCCAACATCTCCCTCCACATTTGTCAACATTTCAAG 3060
Qy 3061 gggccacagtgacttttctgggcaattttcccgagatgtttacagactgtgagacagcag 3120
Db 3061 GGCCACACAGTGACTTTTCTGGGCATTTTCCAGATGTTTACAGACTGTGAGTACAGCAG 3120
Qy 3121 aaatcttttactagtgctgtgtgtatataataaaacaaatgtaaattctcttttagccc 3180
Db 3121 AAAATCTTTTACTAGTGTGTGTGTATATATAAACAATGTAAATTTCTTCTTTAGCCC 3180
Qy 3181 attttctagactctctgtggaatatttctgtgtgtgtatataatcatgtgtgtat 3240
Db 3181 ATTTTCTAGACTGCTCTGTGGAAATATATTGTGTGTGTATATATCATGTGTGTGTAT 3240
Qy 3241 ggtatgtatgatttaacttaactaaataattgtgcccgagttgtgtccaaagtgcata 3300
Db 3241 GGTATGTATGATTAATCTAATTAATTAATGTGCCCGCAGTGTGCCAAGTGCATA 3300
Qy 3301 gctgagcgtaaaactagtggtgttcatcatgacaaacctgcctcagtcctcatttaacc 3360
Db 3301 GTCTGAGCTAAATCTAGTGATTTTTCATCATGACAAACCTGCCTCAGTCCATTTTAACC 3360
Qy 3361 ttagcaacctctgtcattcataactcttaactcatgttaactcattacaaatggatata 3420
Db 3361 TGTAGCAACCTTCTGCATTCATAAATCTGTAAATCATGTTTACCATTACAAATGGGATATA 3420
Qy 3421 agagcgagcgtgaaagcagatgagctgtggactagcaataagggtttttgtgtgtgt 3480
Db 3421 AGAGCAGCGTGAAGCAGATGAGCTGTGGACTAGCATATAGGTTTGTGTGTGTGT 3480
Qy 3481 tggtttgataaagcagtatatttgggggtcaatattgttctctgtgctggagcaaaagtcatta 3540
Db 3481 TGGTTTGATAAAGCAGTATTTGGGGTCATATTGTTTCTGTGCTGGAGCAAAAGTCATTA 3540
Qy 3541 cactttgaagtattatattgttcttctcctcaattcaatgtgggtgagtaaatggcaggt 3600
Db 3541 CACTTTGAAGTATTATATTGTTCTTATCTCTCAATTCATATGTTGGTGTGATGAATATGCCAGGT 3600
Qy 3601 tgtctgatattcttccagacttcgcagacagatgtgctgataataaattaggtaaagata 3660
Db 3601 TGCTGTATATTTCTTTTACAGACTTCGCCACAGAGATGCTGTATATAAATTAAGTAAGATA 3660
Qy 3661 atttgttggggcattattttagcacaggttaaaataacatcaggtttccagttgcttgaattg 3720
Db 3661 ATTTGTTGGGCCATATTTTAGCAGAGTAAATAAATCAATCATCAGTTCACAGTTCCTGTTGAATTG 3720
Qy 3721 caaggttaagaagtactgccttttctgtgtgttagcagtcgaatctattattccactggcg 3780
Db 3721 CAAGGCTAAGAAGTACTGCTTTTGTGTGTGTAGCAGTCAAAATCTATTATCCACTGGCG 3780
Qy 3781 catcatgacgtgatataatgcctataataaagaacataggtttcacaccattttgttttag 3840
Db 3781 CATCATGACGTGATATATGCTTATATAAAGCCATAGGTTTACACCATTTTGTGTAG 3840
Qy 3841 acaattgtcttttttcaagatgcttttcttcttcatatgaaaaaatgcattttataa 3900

```

```

Db 3841 ACAATTGTCTTTTTCAGAGATCCTTGTGTTCTTTTCATATGAAAAAATGCATTTTATAA 3900
Qy 3901 attcagaaagtcataagattttcgaaggcgtcaacgtgcatttttatttatttgcgaactggttaag 3960
Db 3901 ATTCAAGAAAGTCATAGATTTCTGAAGGCGTCAACGTGCATTTTATTATTATGGACTGTAAG 3960
Qy 3961 taactgtgtgtttactcagcaggaattttccaatttctacatttactacattctttcaaca 4020
Db 3961 TAACTGTGGTTTACTAGCAGGAATATTTCCAATTTTACTCTTTACTACATCTTTTCAACA 4020
Qy 4021 agtaactttgtagaatagccagcagaagcccgagtgagtgagtgagtgagtgagtgagtgagtg 4080
Db 4021 AGTAACTTTGTAGAATAGCCAGAGCCCAAGGCCCTGAGTTGGCAGTGGCCCATAAAGTG 4080
Qy 4081 taaaataaaagtttacagaaacctt 4105
Db 4081 TAAATAAAAGTTTACAGAAACCTT 4105

RESULT 5
LOCUS HUMETN1R HUMETN1R 4101 bp mRNA linear PRI 07-FEB-1999
DEFINITION Human mRNA for endothelin-1 receptor.
ACCESSION D90348
VERSION D90348.1 GI:219649
KEYWORDS endothelin-1 receptor.
SOURCE Human placenta, cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 4101)
AUTHORS Hosoda,K., Nakao,K., Hiroshi-Arai, Suga,S., Ogawa,Y., Mukoyama,M., Shirakami,G., Saito,Y., Nakanishi,S. and Imura,H.
TITLE Cloning and expression of human endothelin-1 receptor cDNA
JOURNAL FEBS Lett. 287 (1-2), 23-26 (1991)
MEDLINE 91348221
COMMENT These data kindly submitted in computer readable form by: Kazuwa Nakao
Second Division, Department of Medicine
Kyoto University School of Medicine
54 Shogoin Kawahara-cho
Sakyo-ku, Kyoto 606
Japan
Phone: 075-751-3170
Fax: 075-771-9452
Location/Qualifiers
1. .4101
/organism="Homo sapiens"
/db_xref="taxon:9606"
481. .1764
/note="endothelin-1 receptor"
/codon_start=1
/protein_id="BAA14359.1"
/db_xref="gi:219650"

FEATURES
source
CDS
BASE COUNT 1136 a 859 c 844 g 1262 t
ORIGIN

Query Match 99.9%; Score 4101; DB 9; Length 4101;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 5 tcgcggcgccctcttcgcgtcccgagtgagtggaagctctggagcttggagagagac 64
Db 1 TCGCGCGCCCTCTTCGCGTCCCGAGAGTGAGTGGAAGGCTCGAGCTTGGAGGAGAGAC 60

```

Qy 65 gggagagacagactggaggcgtgttccctccggagttttcttttctgtcgcagccctcgcg 124
Db 61 GGGGAGCAGACACTGGAGGGCGTGTCTCCGGAGCTTTTCTTTTCGTCGAGCCCTCGCG 120
Qy 125 cgcgcgtacagtcacccgcgtgcttcgaacattgtgagagcggtgagagagcttctc 184
Db 121 CGCGCGTACAGTCATCCCGCTGGTCTGACGATTGTGAGAGCGCGGTGAGAGGCTTCATC 180
Qy 185 catcccccgcggtcgtcgcgcgggatcggggtccagcgacacctccccgggagagcag 244
Db 181 CATCCCAACCGGTCTGTCGCCGGGATTTGGGTCCACGGACACCTCCCGGGGAGAAGCAG 240
Qy 245 tggccaggaagttttctgaagccggggaagctgtgcagcggaagcccccgcgcgcgga 304
Db 241 TGCCACGGAAGTTTCTGAAGCGGGGAAGCTGTGCAGCCGAAGCCGCCGCCGCCGCGGA 300
Qy 305 gcccggaacacgcgcacccctccgcgcacccccacccctcgtttctcgcggttctctgcgc 364
Db 301 GCCCGGACACCGGCCACCCCTCCGCGGCCACCCACCCCTCGCTTTCTCGGCTTCTCTGGC 360
Qy 365 ccaggcgcgcgcgacccgcgcagctgtctgcgcagcgcgcagctccacggtgaaaaaaa 424
Db 361 CCAGCGCGCGCGGACCCGCGACGCTGTCTCGGCACGCCGCGGCTCCACGGTGAARAAAA 420
Qy 425 agtgaaggtgtaaaagcagcacaagtgcaataaagagattttctcctcaaatgtgcctcaag 484
Db 421 AGTGAAGGTGTAAAGCAGCACAAGTGCATTAAGAGATATTTCCTCAAAATTTGCCTCAAG 480
Qy 485 atgaaacccctttgcctcagggaatcccttttggctggcactgggttgatgtgtaacagt 544
Db 481 ATGAAAACCCCTTTGCCTCAGGGCATCCCTTTTGGCTGGCACATGGTGGATGTGTAATCAGT 540
Qy 545 gataatcctgagagatcacagcacaaaatctaaagcaatcagtggaatgttccaccactttt 604
Db 541 GATTAATCTGAGAGATACAGCACAAAATCTAAGCAATCATGTGGATGATTTCACCACTTTT 600
Qy 605 cgtggcacagctcagcttctcctgggtaccactcctcaacccactaaatttggctcacc 664
Db 601 CGTGCACAGAGCTCAGCTTCTCGTTACCACTCATCAACCCACTAATTTGGTCTCAACC 660
Qy 665 agcaatgctcaatgcaacaactatggccacagcagactaaaattacttcagcttccaaa 724
Db 661 AGCAATGGCTCAATGCACAACATATTGGCCACAGCAGATAAAATTACTTTCAGCTTTCAAA 720
Qy 725 tacattaactgtgatactctgtactatttctcgtgggaatggtgggaatgcaact 784
Db 721 TACATTACACTGTGATATCTGTACTATTCTATCTGTTGGAATGGTGGGAATGCAACT 780
Qy 785 ctgctcaggatcattaccagaaacaaatgatatgaggaaatggcccccaacgcgctgatacc 844
Db 781 CTGCTCAGGATCAATTTACCAGAAACAAATGTATGAGGAATGGCCCCCAACGCGCTGATAGCC 840
Qy 845 agcttgccttggagaccttatctatgtggtcattgacttccctcctatcaatgattttaag 904
Db 841 AGTCTGGCCCTGGAGACCTTATCTATGTGTGTCATGTATCTCCCTATCAATGATATTAG 900
Qy 905 ctgctggctggcgctggccttttgcatacaaatgacttggcgatttcttcttgcgaagctg 964
Db 901 CTGCTGGCTGGCGCTGGCCCTTTTGATNATCAATGACTTTGGCGTATTTCTTTGCAAGCTG 960
Qy 965 ttcccccttttgcagaagctcctcgtggggatcacccgtcctcctcaacccctcgtccttagt 1024
Db 961 TTCCCTCTTTTGCACAAGTCTCGTGGGGATCACCGCTCTCAACCTCTGCGCTCTTAGT 1020
Qy 1025 gttgacaggtacagagcggttgcctcctcctggagctgcttcaggggaattgggaattcctttg 1084
Db 1021 GTTGACAGGTACAGACAGTGTGCTCCCTCCCTGGAGTCTGCTGTGAGGGAATTTGGGATTCCTTTG 1080
Qy 1085 gtaactgcattgaaatgtctcccatcctggatccctcctccttctcctcctcctggaattcctgaa 1144
Db 1081 GTAACGTGCATGAAATGTCTCCATCTGGATTCCTCTCTTTATCTGCGCCATTCCTGAA 1140

Qy 1145 gcgattggcttcgcatcgggtaccctttgaatatataggggtgaacagcataaaacctgtatg 1204
Db 1141 GCGATTGGCTTCGTCATGGTACCCTTTGAATATAGGGGTGAACAGCATAAAACCTGTATG 1200
Qy 1205 ctcaatcccatcaaaatcatcaggtctcaccagaatgtaaaaggactgggtgctcttc 1264
Db 1201 CTCATGCCACATCAAAATTCATGGAGTTCTACCAAGATGTAAAGGACTGGTGGCTCTTC 1260
Qy 1265 ggggtctattctctgtatgccttgggtgcactcgtatctctcctacacctcactgactgt 1324
Db 1261 GGGTCTATTCTCTATGCCCTTGGTGTGCACCTCGATCTTCTACACCTCATGACTGT 1320
Qy 1325 gagatgttgaacagaagaatggcagcttgagaattgccctcagtgaaacatcttaagcag 1384
Db 1321 GAGATGTTGAACAGAAGGAATGGCAGCTTGAGAAATGGCCCTCAGTGAACATCTTAAGCAG 1380
Qy 1385 cgtcgaaagtgcaaaaacagtttctcgttgggtgttaatttttgcctcttgcctggttc 1444
Db 1381 CGTCGAGAGTGGCAAAACAGTTTCTGCTTGGTGTAAATTTTCTTCTTCTTGGTCTTC 1440
Qy 1445 cctctcacttaagccgtatatattgaagaaaaactgtgtataacgaaatggacaagaaccca 1504
Db 1441 CCTCTTCATTAAAGCCGTATATTGAAGAAAACTGTGTATAACGAAATGGACAAGAACCGA 1500
Qy 1505 tgtgaattacttagtttcttactcgtcagtgattacatcggtattaaacttggcaaccatg 1564
Db 1501 TGTGAATTTACTTAGTTTCTTACTGCTCATGGATTACATCGGTATTAACTTGGCAACCATG 1560
Qy 1565 aatcatgtataaaaccccatagctctgtatttctgagcaagaataataaaattgtttc 1624
Db 1561 AATTCAATGTATAAACCCCATAGCTGTATTTTGTGACGAAGAATAATTAANAATTTGTTTC 1620
Qy 1625 cagtcagtcctcgtcgtcgtgttaccagtcctcaaaagtctcgtatgacctcgtccccatg 1684
Db 1621 CAGTCATGCCCTCTGCTGCTGTGTATACCAGTCCAAAAGTCTGATGACCTCGTCCGCCATG 1680
Qy 1685 aacgggaacaagcatccagtggaagaacacagatcaaaacacacacacacacagcgagc 1744
Db 1681 AACGGAAACAAGCATCCAGTGGGAAGAACCCAGCATCAAAACACCAACACAGACCCGGAGC 1740
Qy 1745 agcataaggcagcatgaactgaacaccttagaagcactcctcgtcgttactcctccataatc 1804
Db 1741 AGCCATAGGACAGCATGAACCTGACCCCTTAGAAGCACCTCTCGGTACTCCCATATTC 1800
Qy 1805 ctctcgggaaaaaaatcacaggcactgtgactcccggggaatctctctctcgtaccttc 1864
Db 1801 CTCCTCGAGAAAAAATCACAAAGCACTGTGACTCCGGAATCTCTTCTCTGATCCTTC 1860
Qy 1865 ttcccttaattcactcccaaccccaagaagaatgctttcccaaaacccgaaggtagactgg 1924
Db 1861 TTCCCTTAATTCACTCCACACCCCAAGAAATGCTTTCCAAAACCGCAAGGTAGACTGG 1920
Qy 1925 ttatcccccacacacatctacgaatcgtactcttcttaattgacttaattacatctct 1984
Db 1921 TTTATCCACCCACACATCTAGGAATCTGACTCTTCTTAAATGATCTAATTTACATATCT 1980
Qy 1985 gctggtgttatccagcactaaaaatgggtggagctggggggaatgaagactgttaaat 2044
Db 1981 GCGTGTCTTATTCAGCACTAAAAATGGTGGAGCTGGGGAGATGAAGACTGTATAAT 2040
Qy 2045 gaaacagaagagatatttactacttttgcataaataagagctttcagtcacatgctag 2104
Db 2041 GAAACAGAAGGATATTACTACTTTTGCATGAAATAGAGCTTTCAAGTACATGGCTAG 2100
Qy 2105 cttttatggcagctcgtggaatgttcaatgggaactgggtcaccatgaaacttttagagat 2164
Db 2101 CTTTTATGGCAGTTCTGGTGAATGTTCATGGGAATGGTCACTGCTCACTGAACCTTTAGAGAT 2160
Qy 2165 taacgacagaagattttctacttttttaagtgattttttgtccttcagccaaacacaaat 2224
Db 2161 TAACGACAGAAGATTTTCTACTTTTTTAAGTGATTTTTTGTCTCTCAGCCAAACAAATAT 2220
Qy 2225 gggctcaggtcactttttatttgaaatgctcatttggcgccagttatttttaactgcataat 2284

Db	2221	GGGCTCAGGTCAC	TTTTATTGAAATGTCATTTGGTGCCAGTATTTTAACTGCATAAT	2280
Qy	2285	agcctaacaatgattt	gaaacttatttacaatagttgaaaaaaagacaaaata	2344
Db	2281	AGCCTAACATGAT	TATTTGAACTTATTTACACATAGTTTGAAAAAAAAGACAAAAATA	2340
Qy	2345	gtattcagtgagcaatt	tagattagttttccacgtcaactattatttttttaaaccac	2404
Db	2341	GTATTCCAGGTGAGCA	ATTAGATTATTTTCCACGTCACTATTATTTTTTAAACAC	2400
Qy	2405	aaattctaaagtaca	aaaaatactactacagggcccttaaacagctgcgtgatgacacatttg	2464
Db	2401	AAATTTCTAAAGT	TACAACAAATACTACAGGCCCTTAAAGCAGCTGTATGACACATTTG	2460
Qy	2465	gcagttttaatagatg	tactcaaaatatttttagaacgtgattttattttttaaatgg	2524
Db	2461	CGAGTTTTAATAG	ATGTTACTCAAGAAATTTTTTAAAGACTGATTTATTTTTTAAATGG	2520
Qy	2525	tgttttattacaagg	gacctgaaacatgttttgatgttaaatccaagaatgacttca	2584
Db	2521	TGTTTTATTACAAG	GACCTTGAACATGTTTGTATGTTAAATTCAAAAGTAATGCTTCA	2580
Qy	2585	atcagatagttctttt	caaaagttcaatactgctgttttttcoatgtaaatttttgtatgaaaa	2644
Db	2581	ATCAGATAGTTCT	TTTTTTCACAAAGTTCAATACTGTTTTTTCATGTAAATTTGTATGAAAA	2640
Qy	2645	tcaatgtcaagta	caccaaatgtaatgatgtgtcatttaactgcgcgtgagacttcaag	2704
Db	2641	TCAATGTCAAGT	TACCANAATGTAATGTATGTTCATTTAACTCTGCCCTGAGACTTTCAG	2700
Qy	2705	tgacgtgtatataga	agttctaaaaacacacactgaagaaaaagatcgaaatttttcagatga	2764
Db	2701	TGCACGTGTATAGA	AGTCTRAAACACACACCTTAAGAGAAAAAGATCGAATTTTTCAGATGA	2760
Qy	2765	tctggaaaattttcat	cctcagggtatttgytaatagtgacatatatatgtatatatacatatcacc	2824
Db	2761	TTCCGGAAATTTCA	PTTCAGGTATTTTGTAAATAGTGACATATATGTATATACATATCAC	2820
Qy	2825	tctattctcttaatt	tttgtttaaagttaactggcagtaagtttttttgtatcattcc	2884
Db	2821	TCCATTTCTCTCT	TAATTTTGTAAATGTTAACTGGCAGTAGTCTTTTTTTCATCATTC	2880
Qy	2885	cttttccatatagga	aaacataattttggaagtgccagatgagtttatcatgtcagtgaaa	2944
Db	2881	CTTTTCCATATAG	GAACATAAATTTTGAAGTGGCCAGATGAGTTTATCATGTCTAGTGAAA	2940
Qy	2945	aataattaccaca	aaaaatgcaccagtaacttaaacgattcttcaacttttgggggtttcag	3004
Db	2941	AATAAATTACCC	ACAAAATGCCACCAGTAACTTTAAACGATTCTTTCACATCTTGGGGT	3000
Qy	3005	tatgaacctaaact	cccccccccaatctcctcccacattgtcaacatttcaaaagggcc	3064
Db	3001	TATGAACCTACT	CCCCACCCCAACATCTCCTCCACATTTGTACCAATTTCAAAAGGGCC	3060
Qy	3065	cacagtgactttgt	ctgggcattttccagatgtttacagactgtgagtcagacgagaaaa	3124
Db	3061	CACAGTGACTTTT	GCTGGGCATTTTCCAGATGTTTACAGACTGTGAGTACAGCAGAAAA	3120
Qy	3125	tcttttactagtg	tggtgtgcatatatataaaacaattgtaaatttcttttagccattt	3184
Db	3121	TCTTTTACTAGT	GTGTGTATATATAAACAATTTGTAATTTCTTTTAGCCCAATTT	3180
Qy	3185	tcttagactgtct	ctgtggaaatatatttgtgtgtgatatatgcattgtgtgatgta	3244
Db	3181	TTCTPAGACTGT	CTGTGGAAATATATTGTGTGTGTATATGCATGTGTGTATGTGTA	3240
Qy	3245	tgtagtgatttaa	cttaactaaataattgtgtcccgagttgtgccaagtgcatagct	3304
Db	3241	TGTATGGATTTA	ATCTAATTAATTTGTGCCCGCAGTTGTGCCAAGTGTCATAGTCT	3300
Qy	3305	gagctaaaaat	ctaggtgtgttcatcatgatgacaacctgcctcagtcctattttaaactgta	3364

3301	GAGCTAAATCTTAGGTGATGTTGTCATCATGACAACGCTGCTCAGTCCATTTTAACTGTA	3360
3365	gcaacctctgcattcataaaactcttgtaatacatggtaccattacaaaagggaataaag	3424
3361	GC AACCTCTG CATTCATAAAATCTTGTATATCATGTTACCATTTACAAATGGGATATAAGAG	3420
3425	gcagcgtgaaagcagatgagctgctggactagcaatataggggttttggttggttggt	3484
3421	GCAGCGTGAAACGACATGAGCTGTGGACTAGCAATATAGGGTTTTGTTGGTTGGT	3480
3485	ttgataaagcagattttgggggtcatattgttccgtgctgagcaaaagtcattcacact	3544
3481	TTGATAAAAGCAGTATTGGGGTCATATGTTTCCGTGCTGGAGCAAAAGTCATTACACT	3540
3545	ttgaagtattatatgtttcttattatcctcaatcgaattgctgataataaaataagtaagataaatt	3604
3541	TTGAAGTATTATATGTTCTTATCCTCAATTCAATGTGGTGATGAAATGGCAGGTTGTC	3600
3605	tgatatctttccagacitccgacagagattgctgataataaaataagtaagataaatt	3664
3601	TGATATTCTTTTCAGACTGCGCACAGATTTGCTGATAATAAAATAGGTAAAGTAAATTT	3660
3665	gttggggccatatatttagcacaggttaaaataacatcaggtctccagttgcttgaattgcaag	3724
3661	GTTGGGCCATATTTTAGCACAGTAAATATACATCAGGTTCCAGTTGCTTGAATTGCAAG	3720
3725	gctaagaagtaactgccttttggttgtagcagtcacaaatctattattccactggcgcatc	3784
3721	GCTAAGAAGTACTGCCCTTTTGTGTGTTAGCAGTCAANTCTATATTCCATGCGGCATC	3780
3785	atatgcagtgatatatgcctataataagaacctaggttcacacatttttttagacaa	3844
3781	ATATGCAGTGATATATGCTATAATATAAGCCATAGGTTACACCATTTTGTTAGACAA	3840
3845	ttgtcttttttcoaagatgctttgtttcttccatagaaaaaatgaatttataaattc	3904
3841	TGTGCTTTTTCAGATGCTTGTGTTCTTTCATATGAAAAAATGCATTTTATAAATTC	3900
3905	agaaagtcatagatcttgaaggcgtaaacgtgaattttatttgactgctgtaagtaac	3964
3901	AGAAAGTCATAGATTTCTGAAGGGCGTCAACGTGCATTTTATTATGGACTGGTAGTAAC	3960
3965	tgtggtttactagcaggaaattttcccaatttctacctttactacatcttttcaacaagta	4024
3961	TGTGGTTTACTAGCAGGAATATTTCCAATTTCTACCTTTACTACATCTTTTCACAAGTA	4020
4025	actttgtagaatatgagccagagccaaggccctgagttggcagtgggcccataagtgtaa	4084
4021	ACTTTGTAGAAATGAGCCAGAGGCCAGGCCCTGAGTTGGCAGTGGCCCACTAAGTGA	4080
4085	ataaaagtttcagaaacctt	4105
4081	ATAAAAGTTTACAGAAACCTT	4101

RESULT	6
S45956	
LOCUS	
DEFINITION	S45956 3305 bp mRNA linear PRI 08-MAY-1993
ACCESSION	endothelin receptor subtype A [human, placenta, mRNA, 3305 nt].
KEYWORDS	S45956
VERSION	S45956.1 GI:257375
SOURCE	human placenta.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 3305)
AUTHORS	Hayzer,D.J., Rose,P.M., Lynch,J.S., Webb,M.L., Kienle,B.K., Liu,E.C., Bogosian,E.A., Brinson,E. and Runge,M.S.
TITLE	Cloning and expression of a human endothelin receptor: subtype A
JOURNAL	Am. J. Med. Sci. 304 (4), 231-238 (1992)
MEDLINE	93035452
REMARK	GenBank staff at the National Library of Medicine created this

Qy	2002	ctaaaaatggtgggagctg999gaaatgaagactgttaatatgaaccagaagatatt	2061
Db	1740	CTAAAAATGGTGGGAGCTGGGGAGAAATGAAGACTGTAAATGAAACACAGAAGATATT	1799
Qy	2062	tactactttgcatgaatatagctcttcaagctacatgctagctctttatgtgcaagtctg	2121
Db	1800	TACTACTTTTGCATGAAATATAGAGCTTTCAGTAGCATGGCTAGCTTTTATGGCAGTCCGT	1859
Qy	2122	gtgaatgttcaatgggaactggtgcaccatgaaacttttagagattaaacgacaagattttct	2181
Db	1860	GTGAATCTTCACGTGGCACTGGTCACCATGAACTTTTAGAGATTAAACGACAAGATTTCT	1919
Qy	2182	actttttttaagtga-ttttttgtcttcagoccaaacacaatatagggtcagtgcaacttt	2240
Db	1920	ACTTTTTTTTAAAGTGATTTTTTTTGGCCTTCAGCCAAACACAATATGGGCTCAAGTCACTTT	1979
Qy	2241	tattgaaatgtcatttggctgcagatattttttaaactgcataaataagcctaactgattat	2300
Db	1980	TATTTGNAATGCTATTTGGTGGCAGTATTTTTTAACTGCATANTAGCCTAACATGATAT	2039
Qy	2301	ttagaacttattcacacatagtttgaaaaaaaagacaaaaatagtattcaggtgagcaa	2360
Db	2040	TTGAACCTTATTTACACATAGTTTGAATAAAAAAAGACAAAAATAGTATTACAGGTGACAA	2099
Qy	2361	ttgagattagtattttccacogtccactatttttttttaaaacacaaattcttaagctaca	2420
Db	2100	TTAGATTAGTATTTTCCACGTCACCTGTTTATTTTTTTTAAAAACAAAAATCTTAAGCTACA	2159
Qy	2421	acaaatctacaggcccttaaaagcagctcgtagacacatttggcagtttaatatagatgt	2480
Db	2160	ACAAATACTACAGGCCCTTAAAGCAGCTGTATGACACATTTGGCAGTTTAAATAGATGT	2219
Qy	2481	tactcaagaattttttaagaactgtatttatttttttaaagtgtttttattacaagg	2540
Db	2220	TACTCAAGAAATTTTTTAAAGACTGTATTTTATTTTTTAAATGCTGTTTATTACAAGGG	2279
Qy	2541	accttgaacatgttttgtatgttaaattcaaaagtaagtcttaactcagatagttctttt	2600
Db	2280	ACCTTGAACATGTTTGTATGTAAATTTCAAAGTAATGCTTCAATCAGATAGTCTTTTT	2339
Qy	2601	tcaasagttccaatcactgttttccatgtaaaatttgtatgaaaaatcaatgtcgaagtacca	2660
Db	2340	TCACAGTTTCAAT-CTGTGTTTTCATGTAAATTCGTATGAAAAATCAATGTCAAGTACCA	2398
Qy	2661	aaatgtaatgtatgttgcatttaactgcctgagacatttcagtcacgtatataagaa	2720
Db	2399	AAATGTTAAATGTATGTGTCATTTAACTCTGCCTGAGACTTTTCAGACACTGTATATAGAA	2458
Qy	2721	gtctaataacacacctaaagaaaaagatcgaaatttttcagatgattcoggaaattttcatt	2780
Db	2459	GTCTAAACACACCTTAAAGAG-AAAAAGATGCGAATTTTTTCAGATGATTCAGAAAAATTTTCATT	2517
Qy	2781	caggtatttggtaatagtgacatatatgtatatcacatcacctcctattctcttaatt	2840
Db	2518	CAGGTATTTGGAATAGTGACATATATATGTATATACATATCACCTCCTCTTCCTCTTAATT	2577
Qy	2841	tttgttaaaatgttaactggcagtaagtcttttttttgatcatctcccttttccatataggaa	2900
Db	2578	TTTTCTTTAAAAATGTTTAACTGGCAGTAAGTCTTTTTTTTGTATCATTCCTTTTTCCATATAGGAA	2637
Qy	2901	acataatttgaatggccagatgatttatacatgtcagtgaaaaataattaccacaaa	2960
Db	2638	ACATAAATTTTGAAGTGGCCAGATGAGTTTATCATGTGAGTGAATAAATATCCACAAA	2697
Qy	2961	tgcaccagtaacttaacgattcttccactctcttggggtttttcagtagaaccttaactccc	3020
Db	2698	TGCCACCAG--AACTTACGATTCTTCACTTCTTGGGGTTTTTCAGTAGAACCTTAACCTCCC	2755
Qy	3021	cacccacaactctccctcccacattgtcaccaatttcaaaagggcccacagtgaattttgct	3080
Db	2756	CACCCCAACATCTCCCTCCACATTTGCACATTTTCAAAGGGCCACAGTGACTTTTGTCT	2815

QY	3081	gggcattttccacagatgttcaagactgtgagtaacagcagaatacttttactagtgtgt	3140
Db	2816	GGGCATTTTCCAGATGTTTTACAGACTGTGAGTACAGCAGAAAACTTTTACTAGTGTT	2875
QY	3141	gtgttatatatataaacaaatgtaaaattcttttttagccattttctagactgtctctg	3200
Db	2876	GFGTGTATATATATAAACAATTGTAAATTTCTTTTAGCCCCATTTTTCFAGACTGTGTC	2935
QY	3201	tgggaataatttggtgtgtgatcatatgctgtgtgtgatgtgtatggtatggtattaact	3260
Db	2936	TGGAATANATATGTGTGTGATATATCATCATGTGTGTGATGATGATGATGATTAANCT	2995
QY	3261	aatctaataattgycgcccgagttgtgccaagtgcatagtctgagctaaaaactagggt	3320
Db	2996	AATCTAAATAATTGCCCGCAGTTGTGCCAAAGTGCATAGTCTGAG-TAAAATCTAGGT	3054
QY	3321	gatgttcacatgacaacctgcctcagtcacattttaacctgtagcaacctctgcattc	3380
Db	3055	GATTGTTTCATCATGACAACCCTGCCTCACTGTCATTTTAACCTGTAGCAACCTTCTGCA	3114
QY	3381	ataaatcttgttaactatgtttaccattucaaatggatataaagagcgagcgtgaaagcaga	3440
Db	3115	ATAATCTTGTAATCACTGTTACCATTACAATGGGATATAAGAGCGGTGAAGGCAGA	3174
QY	3441	tgagctgtggactagcaatatagggtttgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt	3500
Db	3175	TGAGCTGTGGACTAGCAATATAGGGTTTGTGTTGTTGGTTGGTTGGTTGGTTGGTTGG	3234
QY	3501	tggggtcataatgtttctctgtgctggagcagaagtcatcacacttgaagtattattgt	3560
Db	3235	TGGGTCTATATGTTNCTGTGCTGGAGCAAAGTCACTTACACTTTGAAGTATTATATTG	3294
QY	3561	tctttatcct	3570
Db	3295	TTCTTAGCCT	3304
<hr/>			
RESULT	7		
LOCUS	BC022511	2705 bp	mRNA linear PRI 04-FEB-2002
DEFINITION	Homo sapiens, endothelin receptor type A, clone MGC:26548		
ACCESSION	IMAGE:4812050, mRNA, complete cds.		
VERSION	BC022511		
KEYWORDS	BC022511.1 GI:18490297		
SOURCE	MGC.		
ORGANISM	human.		
REFERENCE	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 2705), Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (01-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. cDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305 Web site: http://www-shgc.stanford.edu Contact: (Dickson, Mark) mcd@paxil.stanford.edu Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.		
<hr/>			
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAK Plate: 32 Row: 1 Column: 7			

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503464.

FEATURES
source 1. .2705
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="LocusID:1909"
/db_xref="taxon:9606"
/clone="MGC:26548 IMAGE:4812050"
/tissue_type="Brain, hippocampus"
/clone_lib="NH.MGC_95"
/lab_host="DH10B"
/note="Vector: pBluescript"
351. .1634
CDS
/codon_start=1
/product="endothelin receptor type A"
/protein_id="AAH22511.1"
/db_xref="GI:18490298"
/translation="METLCRLASFALVCCVISDNPERYSTNLSNHVDFTFRTGE
LSFLVTHQPTNLVPSNGSMHNYCPOOTKITSAFKYINTVVISIFIVGMVGNATLL
RIYQNKCMRGNHNLASIALGLDIYVLDLPINVKLLAGRPFDHDFGVFLCKL
FPLQKSSVGITVNLICASVDRYRAVMSRVQIGIPLVTAIEIVSIWILSFILAI
PEALGFVMPPEYRGEOHKTCMLNATSKEMEFYODVKDWLFGFYFCMPLVCTAIFYT
LMTCMLNRNGLSRJALSEHLKORREVAKTVECLVVI FALCHFPVHLRLKTKTVN
EMOKNCELSPFLMDYIGINLATMNSCINPIALFYVSKFKNCFQSCCLCCCYQSK
SLMTSPMNGTSIQWDHNDQNNHNTDRSSHKSMN"
BASE COUNT 767 a .594 c 560 g 784 t
ORIGIN

Query Match 63.68; Score 2609.6; DB 9; Length 2705;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 2673; Conservative 0; Mismatches 14; Indels 5; Gaps 5;
Qy 134 agtcatcccgctgtctgacatgtgagagcggtgagagcgttcattcatccacc 193
Db 1 ACTCATCCCGTGGTCTGACATTGTGGAGAGCGGTGGAGAGGCTTCATCCATCCCACC 60
Qy 194 cggctcgtcgccgggattggggtcccaagcgacacctcccgggagagcagtgcccagga 253
Db 61 CGGTCGTCCGCGGGGATTGGGGTCCCAGCGAGACTCCCGCGGAGAACGAGTGGCCAGGA 120
Qy 254 agttttctgaagcgggaaagtctgagccgaagcgcgcgcgcgcgcgcgcgcgcgcgcgcgc 313
Db 121 AGTTTCTGAAGCGCGGGAAGCTGTGCAGCGAAAGCGCCGCCGCGGAGCGCGGAC 180
Qy 314 accgcccacctccgcgcacccacctcgtcttctcgggtctcctcgtcgcgcgcgcgcgcgcgcgc 373
Db 181 ACCGCCACCTCCCGGCCACCCACCCTCGCGGCTCCGGGTCTCTGGCCCGCCGCC 240
Qy 374 gcgggaccccgagctgtctgcgaecgc 433
Db 241 GCGCGGACCGGACGTCTGCGCACCGCGAGCTCCACGGTG-AAAAAAGTCAAGGT 299
Qy 434 gtaaaagcagacaagtgcataagagataattctcaaaattgcctcaagatggaacc 493
Db 300 GTAAAGCAGCAGAGTGAATAGAGATATTTCCTCAAAATTTGCCTCAAGATGGAACC 359
Qy 494 ctctgctcaggcatccttttgctgacactggttgatgtgtaactcagtgataatcct 553
Db 360 CTTTGCCCTCAGGGATCCTTTGGCTGGCACTGGTGTGATGTGTAATCAGTGAATCCT 419
Qy 554 gagagatacacacaaattctaagaatacatgtgatgatttcaccacttttcgtggcaca 613
Db 420 GAGAGATACACACAAATCTAAGCAATCATGTGGATGATTTCCACCACATTTTCGTGGCACA 479
Qy 614 gagctcagcttctcgttggttaccactcatcaaccactaaatttgggtccaccagcaatggc 673
Db 480 GAGCTCAGCTTCTGGTTTACCACCTCATCAAGCCACTAATTTGGTCTCTACCCAGCAATGGC 539
Qy 674 tcaatgcaactatttgccacagcagactaaataacttcagctttcacaatacattaac 733
Db 540 TCAATGCACAACTATTGGCCACAGCAGACTAAATTAATTACTTACGCTTTCAATATTAAC 599

Qy 734 actgtgatattctgtactatttttcatcgtgggaatggtgggaatgcaactctctcagg 793
Db 600 ACTGTGATATCTTGTACTATTCTTGTGGAAATGTGGGAATGCAACTCTGCTCAGG 659
Qy 794 atcatttaccagaacaaatgtatgagaatgcccccaacgcgcgtgtagccagcttgcc 853
Db 660 ATCATTTACCAGAACAAATGTATGAGGAATGGCCACACGCGCTGATAGCCAGTCTGCC 719
Qy 854 ctggagaccttatctatgtggtcattgatctccctatcaaatgtattttaaactgtcgtgct 913
Db 720 CTTGGAGACCTTATCTATGTGGTGCATGTGATCCCTATCAATGATTTAAGCTGCTGGCT 779
Qy 914 gggcgctggccttttgatcacaaatgacttttggcgattattcttgcagcgtcttccccttt 973
Db 780 GGGCGCTGGCCTTTTGTATCACAATGACTTTGGCGTATTCTTGTGCAAGCTGTCCCTTT 839
Qy 974 ttgcagaagctcctcgggtggggtcacgcgtcctcaacctcgcctcttagtggtagacagg 1033
Db 840 TTGCAGAAAGTCTCGGTGGGATCACCGTCTCAACCTCTCGCTCTTTAGTGTGTGACAGG 899
Qy 1034 tacagagcagttgctcctcgtgagtcgtgttcaggaatgggattccttcttggtaactgcc 1093
Db 900 TACAGAGCAGTTGCTCTCGTGAGTCGTCTTCCAGGGAATTTGGATTTCCCTTTGGTAACGCC 959
Qy 1094 attgaaattgctccatctggtgactcgtcttcttattcctcgtccattcctgaagcgtattg 1153
Db 960 ATTGAAATTTGCTCCATCTGGATCCTGTGCTTATCTCTGGCATTTCTCTGAAGCGATTGGC 1019
Qy 1154 ttctcctggtaccctttgaaataggggtgaacagcagcaataaaactgtatcctaagc 1213
Db 1020 TTCGTCTATGGTACCCTTTGAATATAGGGGTGAACAGCATAAACCTGTATGCTCAANTGCC 1079
Qy 1214 acatcaaaattcatggagttctaccagaatgtaaaaggactggtggtccttcgggttcttat 1273
Db 1080 ACATCAAAATTCATGGAGTTCTACCAAGATGTAAAGGACTGCTGCTCTTCCGGTTCAT 1139
Qy 1274 ttctgtagcctcgtgggtgacgtcgtatcttctacacccctcagactgtgagatgttg 1333
Db 1140 TTCTGTATGCCCTTGGTGTGCACCTCTTCTACACCCCTCATGACTGTGTGAGATGTG 1199
Qy 1334 aacagaaggaatggcagcttgagaattgcctcagtgacatcttaacagcgcgcgcgaa 1393
Db 1200 AACAGAAGGAATGGCAGGTTGAGAAATTCCTCAGTGAACATCTTAACAGCGTGCAGAA 1259
Qy 1394 gtgcaaaaacagtttctgctggtgtgtaatttttgcctcttctgctggttccctctcac 1453
Db 1260 GTGCAAAAACAGTTTCTGCTGTTGTTCTTAATTTTGTCTTCTTGTGCTTCCCTGTTTAT 1319
Qy 1454 ttaagcgtatattgaagaaacactgtgtataacgaaatggacagaacacgagatgaaatta 1513
Db 1320 TTAAGCCGTATATTGAAGAAACTGTGTATAACGAGATGGACAGAACCGATGTGAATTA 1379
Qy 1514 ctgagttcttactgctcatgattacatcggtatctaacttggcaacctgaatcactgt 1573
Db 1380 CTTAGTTTCTTACTGCTCATGATTTACATCGGTATTACTTTGGCAACCATGAAATTCATGT 1439
Qy 1574 ataaaccccatagctcgttattttgtgagcaagaattttaaataattgtttccagtcagtc 1633
Db 1440 ATAAACCCCATAGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTGTTTCCAGTCATGC 1499
Qy 1634 ctctgctgctgctgttaccagttcaaaagtctgtgactcgttccctcagcagcagca 1693
Db 1500 CTCTGCTGCTGCTGTGTACAGTCCAAAAGTCTGTATGACTCGGTGCTCCCATGAACGGAACA 1559
Qy 1694 agcatccagttgagaagaccagatcaaaacaacacacacagacgcgcgcgcgcgcgcgcgcgc 1753
Db 1560 AGCATCCAGTGGAGAGCACCGATCAAAACACACACACAGACCGGAGCAGCCATTAAG 1619
Qy 1754 gacagcatgaactgaccaccttagaagcactcctcgggtactccccataatcctctcggag 1813
Db 1620 CACAGCATGAACCTGACACACCTTAGAAGCACTCTCTCGGTACTCCCATTAATCTCTCGGAG 1679
Qy 1814 aaaaaaatcacaggcaactgtgactccgggaactctctctctctctctctctctctctctta 1873

```
Db 1680 AAAAAAAAAACACAGGCAACTGTGAGTCGGGAATCTCTTCTGATCCTTCTCTTAAT 1739
Qy 1874 tcactccacacccaagaagaatgcttcccaaaacgcgaa-gtagactggtttatcca 1932
Db 1740 TCACCTCCACACCCAGAGAGAAATGCTTCCAAAACCGAAGGGTAGACGGTTATPCCA 1799
Qy 1933 cccacaacatctacgaatcgactctcttaattgatctctaattacatatcttcgctgttg 1992
Db 1800 CCCACAACATCTACGAATCGTACTCTTTAAATGATCTAATTTACATATTTCTCGGTGTG 1859
Qy 1993 tattaagcactaaaaaattggtggagactggggagaaatgaagactgtaaatgaaccag 2052
Db 1860 TATTTCAGCACTAAAAAATGCTGGGAGCTGGGGAGAAATGAAGACTGTGTAATGAACCCAG 1919
Qy 2053 aaggatattactactcttctgcatgaaatagactcttcagttacatgctagcttattatg 2112
Db 1920 AGGATATTTACTACCTTTTGCATGAAATAGAGCTTTCAAGTACATGGCTAGCTTTATG 1979
Qy 2113 cgagttctggtgaatttcaatgggaactgggtcaccatgaaactttagagattaaacgaca 2172
Db 1980 GCAGTTCTGGTGAATGTTCAATGGGAACGGTGCACCATGAACTTTAGAGATTAACGACA 2039
Qy 2173 agattttactcttttttaagtga-tttttgtccttcagcccaaacacataatgggctca 2231
Db 2040 AGATTTTCTACTTTTTTTTAAAGTGATTTTTTGTGCTTCAGCCAAACACATAATGGCTCA 2099
Qy 2232 gtcacttttttgaatgctttggtgccaagtatttttaactgcataatagcctaa 2291
Db 2100 AGTCACTTTTTATTTGAATGTCATTTGGTGCCAGTATTTTTTAACTCGCAATAATAGCCTAA 2159
Qy 2292 catgattatttgaacttattcacatagtttg-aaaaaagaaagcaaaaatagttatc 2350
Db 2160 CATGATTATTGAACCTATTACACATAGTTTGAAAAAAGAAAAAGACAAAAATAGTATTC 2219
Qy 2351 aggtgagcaattagtagtattttccacgctactatttttttttaaaacacaattc 2410
Db 2220 AGGTGAGCAATTAGATTAGTATTTCACGTCCTACTGTTTATTTTTTAAACACACAATTC 2279
Qy 2411 taaagctacaacaatactacagggcccttaagacagctgtagcacatttggcagtt 2470
Db 2280 TAAAGCTACAACAATACTACAGGCCCTTAAAGCACAGTCTGATGACACATTTGGCAGTT 2339
Qy 2471 taatagatgttactcaagaattttttaagaactgtatttttttttaaatggtgtttt 2530
Db 2340 TAATAGATGTTACTCAAGAAATTTTAAAGACTGATTTTATTTTTTAAATGGTGTTT 2399
Qy 2531 attacaagggacctggaacatgtttgtatgttaattcaaaagtaagtcttcaatcaga 2590
Db 2400 ATTACAAGGGACCTTGAACATGTTTGTATGTTAAATTCAAAAGTAATGCTTCAATCAGA 2459
Qy 2591 tagttcttttcacaagttcaatactgtttttcatgtaaatatttgcatagaataatcagt 2650
Db 2460 TAGTTCTTTTTCACAAGTTCAAT-CTGTTTTTCATGTAATTTTGTATGAAAAATCAATG 2518
Qy 2651 tcaagtcacaaaattgttaagtgtgtgtcatttaactgtcgtgagactttcagtgcact 2710
Db 2519 TCAAGTACCAAAATGTTAATGATGTGTCACTTAACCTGCTGAGACTTTTCAAGTGCATC 2578
Qy 2711 gtatagaagctcaaaacacacacacacacacacacacacacacacacacacacacacac 2770
Db 2579 GTATAGAAGTCTAAAAACACACACCTTAAGAGAAAAAGATCGAATTTTTCAGATGATTCAGA 2638
Qy 2771 aatttcattcaggtatttgaatagtgacatatatatgtatatcatatca 2822
Db 2639 AATTTTCATTGAGTATTGTTAATAGTGACATATATATATATATATATATATATATCA 2690

RESULT 8
HUMETAR8
LOCUS HUMETAR8 2595 bp DNA linear PRI 14-APR-2000
DEFINITION Human DNA for endothelin-A receptor, exon 8 and 3', flanking region.
ACCESSION D11151
```

```
VERSION D11151.1 GI:219628
KEYWORDS G protein-coupled receptor; endothelin; endothelin-A receptor;
endothelium; smooth muscle.
SEGMENT 8 of 8
SOURCE Homo sapiens DNA, clone_lib:lambda EMBL3 and lambda EMBL3SP6/T7.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2595)
AUTHORS Hosoda,K., Nakao,K., Tamura,N., Arai,H., Ogawa,Y., Suga,S.,
Nakanishi,S. and Imura,H.
TITLE Organization, structure, chromosomal assignment, and expression of
the gene encoding the human endothelin-A receptor
J. Biol. Chem. 267 (26), 18797-18804 (1992)
JOURNAL 92406798
MEDLINE
REFERENCE 2 (bases 1 to 2595)
AUTHORS Hosoda,K.
TITLE Direct Submission
JOURNAL Submitted (18-MAY-1992) Kiminori Hosoda, Kyoto University School of
Medicine, Department of Medicine: 54 Shogoin Kawahara-cho,
Sakyo-ku, Kyoto, Kyoto 606, Japan (Tel:075-751-3170,
Fax:075-771-9452)
COMMENT Submitted (18-May-1992) to DDBJ by:
Kiminori Hosoda
Second Division
Department of Medicine
Kyoto University School of Medicine
54 Shogoin, Kawahara-cho
Sakyo-ku
Kyoto 606
Japan
Phone: 075-751-3170
Fax: 075-771-9452.
FEATURES
Source
1. .2595
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="lambda EMBL3 and lambda EMBL3SP6/T7"
1010(D11145.1:92 511,D11146.1:21. 148,D11147.1:21. 219,
D11148.1:21. 173,D11149.1:21. 154,D11150.1:21. 129,
19. 159)
/codon_start=1
/product="endothelin-A receptor"
/protein_id="BAA01920.1"
/db_xref="GI:219630"
/translation="METICLRASFALVGCVISDNPERKSTNLSNHVDDFTFRGTE
LSFLVTHQPTNLVLPNSNGSMNYCPQTKITSFAKYINTVISCTIFIVGMVGNATLL
RIYQNKCMRNGFNALIASLALGDLIVYVDLPINVKLAGRPFDHNDGVLCKL
FPFLQKSSVGTITVLNICALSVDRYRAVASRSVQIGIPLVTAIEIVSIWILSFILAI
PEAIGVMVPEYRGEQHKTCMLNATSKEMEFYQDVKDWLFGYFCMPLVCTAIFT
LMTCEMLNRRNGSLRLALSEHLKORREVAKTVCFLVIVFALCFPHLSRLKKTYN
EMDKNRCELLSFLLMDYIGINLATMNSCINPIALYFVSKFKNCFQSCLCGCCYQSK
SLMTSVPMNGTSTQWKNHNDNRSHKDSMN"
intron <1..18
exon /number=7
19..2495
polyA_signal /number=8
2474..2479
BASE COUNT 788 a 478 c 459 g 870 t
ORIGIN Chromosome 4.
Query Match 60.2%; Score 2470; DB 9; Length 2595;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2481; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 1624 ccagtcattcctctgctgctgttaccagtcacaaagctcagtcctccccat 1683
Db 15 CCAGTCATCCCTCTGCTGCTGTGTACAGTCCAAAAGTCTGATGACCTCGGTCCCAT 74
Qy 1684 gaacgacaagaagcattccagtgaagaaccacacacacacacacacacacacacacac 1743
Db 75 GAACGGAACAAGCATCCAGTGGGAAGAACCCAGCATCAAAACACACACACACACACCGGAG 134
```


Qy 1744 cagccataagagacagcagtaagtaagcacccttagaagcactcctcggaactcccatataat 1803
Dy 135 CAGCCATAAAGGACAGCATGAAGTACACCAACCCCTTGAAGCACCTCGGTACTCCCATAAAT 194
Qy 1804 cctcgcgagagaaaaaacacagcagcactgtagctcggaatctctctctgactcctt 1863
Dy 195 CCTCTCGGAGAGAAAAATACAGAGCAACTGTGACTCGGGAATCTCTTCTGATCCTT 254
Qy 1864 ctctcttaattcaactccacaccccaagaagaatgctttccaaacccgcaaggtagactg 1923
Dy 255 CTCTCTTAATTCACTCCACACCCCAAGAGAAATGCTTTCCAAACCCCAAGGTAGACTG 314
Qy 1924 gtttatccaccccaaacatctacgaatcgtagctctctttaaattgatctaatattaatc 1983
Dy 315 GTTTATCCACCCCAACAATCTACGAATCGTACTCTTTAAATTGATCTAAATTACATATTC 374
Qy 1984 tgcgtgtgtattcagcactaaaaatggtggagctggggagagaaatgaagactgttaaa 2043
Dy 375 TCGGTGTGTATTCAGCACTAAAAAATGGTGGGAGCTGGGGAGAAATGAAGACTGTTAAA 434
Qy 2044 tgaacccagaaggattattactacttttgcatagaataagagctttccaagtacatggcta 2103
Dy 435 TGAACCAAGAGGATATTTACTACTTTTGCATGAAATAGAGCTTTTCAAGTACATGGCTA 494
Qy 2104 gcttttatggcagttctggtggaatgctcaatgggaactgggtcaccatgaacttttagaga 2163
Dy 495 GCTTTTATGGCAGTCTGTGGTGAATGTTCAATGGGAACCTGGTCACCATGAAACTTTTAGAGA 554
Qy 2164 ttaacgcagaagatttctactctttttaaagtatttttgcctcagcccaaacacaata 2223
Dy 555 TTAACGCACAAGATTTCTACTTTTAAAGTGATTTTTTGTCTTCACGCCAAACACAATA 614
Qy 2224 tgggctcaggtcacttttatttgaatgctatcttgggtgccagttttttaaactgcaata 2283
Dy 615 TGGGCTCAGGTCACCTTTTATTTGAAATGCTATTTGGTGCCAGTATTTTAACTGCATAA 674
Qy 2284 tagcctaacatgattattgaacttattacacatagtttgaaaaaaagagacaaaaat 2343
Dy 675 TAGCCTAACATGATTTATTTGAACCTATTTTACACATAGTTTGAATAAAAAAGACAAAAAT 734
Qy 2344 agtattcaggtgagcaaatagattagttatttccacgtcactatttattttttaaaca 2403
Dy 735 AGTATTCAGGTGAGCAATAGATAGTATTTTCCACGTCACATTTATTTTAAAAACA 794
Qy 2404 caaattcaaaagctacaacaaactacagcgcccttaaagcacagctctgatgacacattt 2463
Dy 795 CAATTTCTTAAGCTACACAATACTACAGGCCCTTAAGCACAGCTGTGATGACACATTT 854
Qy 2464 ggcagtttaataagatggttactcaaaagaattttttaagaactgtagtttatttttaaatg 2523
Dy 855 GGCAGTTTAAATAGATGTTTACTCAAGAATTTTTTAAAGAACTGTATTTTATTTTAAATG 914
Qy 2524 gtgttttattacaaggaccttgaaactgtttgtatgttaaatcaaatcaaatgaacttc 2583
Dy 915 GTGTTTTATTAACAGGGACCTTGAACATGTTTGTATGTATAATTTCAAAAGTAATGCTTC 974
Qy 2584 aatcagatagttctttttccaaagtccaactggtttttcatgtaaattttgtatgaaaa 2643
Dy 975 AATCAGATAGTCTTTTTCACAAGTTCAAT-CTGTTTTTCATGTAAATTTTGTATGANA 1033
Qy 2644 atcaatgtcaagttaccaaaatgttaattgtatgtgtcatcttaactctgctgagactttca 2703
Dy 1034 ATCAATGTCAAGTACCAAAATGTTAATGTATGTGTCAATTAACCTGTGCTGAGACTTTCA 1093
Qy 2704 gtgaactatataagaactcaaacacacacataagagaaaaagatcgaatttttcaagatg 2763
Dy 1094 GTGCAGCTGTATAGAAGCTCAAAACACACACCTTAAGAGAAAAAGATCGAATTTTTCAGATG 1153
Qy 2764 attcggaaattttcattcaggtatttgaatagtagacatatatatgtatatcatatcac 2823
Dy 1154 ATTCGGAATTTTCAATTGAGGTATTTGTAAATAGTGACATATATATGTATATACATATCAC 1213

Qy 2824 ctctatttcttctaattttttttaaaatgttaactggcagtaagtcttttttgatcattc 2883
Dy 1214 CTCCTATTTCTCTTAATTTTGTAAAAATGTTAACTGGCAGTAAGTCTTTTTTGATCATTC 1273
Qy 2884 cctttccatagagaaacaataatttgaagtggccagatgagtttatacatgtcagtgaa 2943
Dy 1274 CCTTTTCCATATAGAAACATAATTTTGAAGTGGCCAGATGATTTATCATGTCTAGTGAA 1333
Qy 2944 aataattaccocccaaaaatgcccacagtaacttaacgaattcttcaactcttctggggttttca 3003
Dy 1334 AAATAATTTACCACAAAATGCCACCAGTAACCTTAACGATTTCTACTTCTTGGGGTTTCA 1393
Qy 3004 gtatgaaccttaactcccccaccccaacatctccctcccacattgccaacttttcaaaagcc 3063
Dy 1394 GTATGAACCTTAACCTCCCCACCCCAACATCTCCCTCCCACATTTGTCAACCATTTCAANGGC 1453
Qy 3064 ccaagtgacttttgcctgggcattttcccagatgtttcacagactgtagtagcacagaaaa 3123
Dy 1454 CCACAGTGACTTTTGTCTGGGCATTTTCCAGATGTTTACAGACTGTGAGTACAGCAGAAA 1513
Qy 3124 atctttactagtgtgtgtgtatataataaaacaatgttaaaattctcttttagccatt 3183
Dy 1514 ATCTTTTACTAGTGTGTGTGTATATATAAAACAATTTGAAATTTCTTTTAGCCCATTT 1573
Qy 3184 tttctgaactgtctctgtggaatatattgtgtgtgatatatgatcatgtgtgtatggt 3243
Dy 1574 TTTCTAGACTGTCTCTGGGAATATATTTGTGTGTGATATATGATGATGATGATGAT 1633
Qy 3244 atglatggatttaacttaacttaataattgtgcccgagttgtgccaaaagtgcatagtc 3303
Dy 1634 ATGTATGGATTTAATCTAATCTAATAATTTGTGCCCGCAGTTGTGCCAAAGTGCATAGTC 1693
Qy 3304 tgagctaaaaactctagggtgattgttcaatcatgacaacctgcctcagctccattttaacctgt 3363
Dy 1694 TGAGCTAAAACTTAGTGTATTTGTCATCATGACAACCTGCGCTCAGTCCATTTTAAACCTGT 1753
Qy 3364 agcaacctcttgcattcaaaaacttgaatcatgttaccattaccataaagatggatataaga 3423
Dy 1754 AGCAACCTTTCGCATTTAATAATCTGTAAATCATGTTTACCATTTACAAATGGGATATAAGA 1813
Qy 3424 ggcagcgtgaaagcagatgagctgtggactagcaataaggggttttctgtgtgtgtgtgtgt 3483
Dy 1814 GGCAGCGTAAAGCAGATGAGCTGTGGACTAGCAATATAGGGTTTTTGTGTGTGTGTGTGG 1873
Qy 3484 tttgataaagcagttatttggggtcataattgtttcctgtgctggagcaaaagtcatcacac 3543
Dy 1874 TTTGATAAAGCAGTATTTGGGGTGCATATTTGTTCTGCTGGAGCAAAAGTCATTACAC 1933
Qy 3544 tttgaagtattatattgttttcttctcaatcaattcaattgggtgataaattgcccaggttgt 3603
Dy 1934 TTTGAAGTATTATATTTGTTCTTATCCTCAATTTCAATTTGGGTGATGAAATTTGCCAGGTGT 1993
Qy 3604 ctgattcttcttcagactctgcagacagatgtgctgataataaaattaggtaagaataatt 3663
Dy 1994 CTGATATTTCTTTTCAGACTTCCCCAGACAGATTTGCTGATATAATAATTTAGGTAAAGATAAT 2053
Qy 3664 tgttggggccatatttttaggacaggttaaaataacatcaggtttccagttgcttgaattgcaa 3723
Dy 2054 TGTGGGCCATATTTTAGGACAGGTAATAATAACATCAGGTTCCAGTTGCTTGAATTGCAA 2113
Qy 3724 ggctaagaagtactgccccttttgtgtgttagcagtcgaacttatatttccactggcgcat 3783
Dy 2114 GGCTAAGAAGTACTGCCCTTTTGTGTGTAGCAGTCAAACTTATTTATTTTCCACTGGCGGAT 2173
Qy 3784 catatgcagtgatataatgcctataataaagcatalaggttcacacacattttgttttagaca 3843
Dy 2174 CATATGCGAGTGNATATATGCCTATAATAATATAGCCATAGGTTTCACACCATTTTGTAGACA 2233
Qy 3844 attgtcttttttttcaagatgctttgtttcttcttcatatgaaaaaaatgcattttataaatt 3903
Dy 2234 ATTGTCTTTTTTCAAGATGCTTTTGTCTTTCATATGAAAAAATGCAATTTTATAAATTT 2293
Qy 3904 cagaaagtcataagatttctgaaggcgtaacgtgcattttatttatggactggtaagtaa 3963


```

|||||
Db 2294 CAGAAAGTCATAGATTCTGAAGCGCTACAGTCGCAATTTATTATTGAGCTGCTAGTAA 2353
QY 3964 ctgtggtttactagcaggaaatttccaatttctactttactacattctttcaacaagt 4023
Db 2354 CTTGTGTTTACTAGCAGGATATTTTCATTTCTACTTTTACTTACATCTTTTCAACAAGT 2413
QY 4024 aacttgtgaaatgatgccagaagcccaagccctgagttggcagtgcccaatgaagttaa 4083
Db 2414 AACTTTGTAGAAATGAGCCAGAGCCCAAGGCCCTGAGTTGGCAGTGGCCCAATAGTGTA 2473
QY 4084 aataaagtttacagaacctt 4105
Db 2474 AATAAAGTTTACAGAACTT 2495
|||||
RESULT 9
AC093908
LOCUS Homo sapiens BAC clone RP11-752L20 from 4, complete sequence.
ACCESSION AC093908
VERSION AC093908.3
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Sulston, J.E. and Waterston, R.
TITLE Toward a complete human genome sequence
JOURNAL Genome Res. 8 (11), 1097-1108 (1998)
MEDLINE 99063792
REFERENCE 2 (bases 1 to 164920)
AUTHORS Paulson, E., Cotton, M. and Creason, K.
TITLE The sequence of Homo sapiens BAC clone RP11-752L20
JOURNAL Unpublished (2001)
REFERENCE 3 (bases 1 to 164920)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (10-SEP-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 4 (bases 1 to 164920)
AUTHORS Waterston, R.H.
TITLE Direct Submission
JOURNAL Submitted (05-FEB-2002) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
REFERENCE 5 (bases 1 to 164920)
AUTHORS Waterston, R.
TITLE Direct Submission
JOURNAL Submitted (08-FEB-2002) Department of Genetics, Washington
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
On Feb 5, 2002 this sequence version replaced gi:15778805.
-----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc
Contact: sapiens@watson.wustl.edu
Center project name: H_NH0752L20
Drafting Center: WIBR
-----

```

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such

as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Woon, P. Y., Zhao, B., Frengen, E., Tateno, M., Catanese, J. J. and de Jong, P. J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at the Roswell Park Cancer Institute (<http://baopac.med.buffalo.edu>)

VECTOR: pBACE3.6

NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the right is RP11-56F3. Actual start of this clone is at base position 1 of RP11-752L20; actual end is at base position 164920 of RP11-752L20.

Data from AC083898 was used to finish this clone, AC093908.

The sequence of AC067873 has been incorporated into AC093908.

FEATURES	Location/Qualifiers
source	1..164920
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="4"
	/map="4"
	/clone="RP11-752L20"
	/clone_lib="RPCI-11"
	3..112
repeat_region	/rpt_family="Alu"
repeat_region	323..393
	/rpt_family="L1"
repeat_region	4328..4359
	/rpt_family="AT-rich"
repeat_region	4786..4811
	/rpt_family="(CA)n"
repeat_region	4813..4874
	/rpt_family="(TC)n"
repeat_region	4883..4910
	/rpt_family="(CA)n"
repeat_region	4948..5511
	/rpt_family="L1"
repeat_region	5058..5099
	/rpt_family="AT-rich"
repeat_region	5129..5164
	/rpt_family="AT-rich"
misc_feature	5149..5154
	/note="similar to Homo sapiens EST AW150086 (NID:96198078)"
	xg36f05.x1"
repeat_region	5837..5927
	/rpt_family="MERL_type"
repeat_region	5930..6038
	/rpt_family="L2"
repeat_region	6126..6196
	/rpt_family="MERL_type"
repeat_region	6261..7868
	/rpt_family="L2"
repeat_region	6671..6698
	/rpt_family="(TC)n"
misc_feature	7297..7299
	/note="similar to Homo sapiens EST BE857602 (NID:gi0371791)"
repeat_region	7694..8060
	/rpt_family="ERV1"

Db 94377 AAATCAATGTCAGTACCAAAATGTAATGTATGTGTGCTAATTTAACTCTGCCTGAGACTTT 94436
Qy 2702 cagtgcactgatatagaaagctctaaacacacacacactaaagagaaagatgaaatttttcaga 2761
Db 94437 CAGTGCACGTATATAGAAAGCTAAACACACACCTAAGAGAGAAAAGATCGAATTTTCAGA 94496
Qy 2762 tgattcggaattttccatcctcaggtatttgtaatatagacatatatatatatacatatc 2821
Db 94497 TCAATTCAGAAAATTTTCATTCAGGTATTTGTAATAGTAGACATATATATATATACATATC 94556
Qy 2822 acctcctattcttaattttttgtaaaatgttaactggcagtaagtattttttgatcat 2881
Db 94557 ACCTCCTATTCTCTTAATTTTCTTAAATGTTAACTGGCAGTAAGTCTTTTTCATCAT 94616
Qy 2882 tcccttttccatataggaaacataattttgaagtggccagatgagttttcatcatgctcagt 2941
Db 94617 TCCCTTTTCCATATAGGAAACATAATTTTGAAGTGGCCAGATGAGTTTATCATGTCAGTG 94676
Qy 2942 aaaaataattaccacacaaatgccaccagtaacttaacagattcttccacttcttggggtttt 3001
Db 94677 AAAAATAATTTACCCACAAATGCCACCAG- -AACTTACGATTCTTCACCTCTTGGGGTTT 94734
Qy 3002 cagatgaacctaaactcccaccccaacacatctccctcccacattgtccaccatttcaagg 3061
Db 94735 CAGTATGAACCTAACTCCCCACCCCAACATCTCCCTCCACATTGTCCACATTTCAAAGG 94794
Qy 3062 gccacacagtacttttgcggcattttccacagattttcagacgtgagttacagcaga 3121
Db 94795 GCCACAGTGACTTTTGGTGGGCAATTTCCACAGATGTTTACAGACTGTGAGTACAGCAGA 94854
Qy 3122 aaatttttactagtggtgtgtatatataataaacaattgttaattcttttagccca 3181
Db 94855 AAATCTTTTACTAGTGTGTGTGTATATATAAACAATTTGTAATTTCTTTTAGCCCA 94914
Qy 3182 ttttctagactgctctgtggaatatattgtgtgtgtgatatatgcatgtgtgtatg 3241
Db 94915 TTTTTCFAGACTGTCTCTGTGGAAATATATTGTGTGTGTATATGATGATGTGTGTGATG 94974
Qy 3242 gtatgtatggaatttaacttaactaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3301
Db 94975 GTATGTATGATTTAATCTAATCTAATTAATTTGTGCCCCGAGTGTGCCAAAGTGATAG 95034
Qy 3302 tctgagtaaaatcagtggt 3361
Db 95035 TCTGAG- TAAATCTAGGTGATTTCTCATCATGACAACTTCCCTCAGTCCATTTTAACTT 95093
Qy 3362 gtagcaaccttctgcatcataaattctgttaattcatgtttaccattacaaatgggatataa 3421
Db 95094 GTAGCAACCTTCTGCATTCATAATCTTGTAAATCTTGTAAATCATGTATTACCAATACAAATGGGATATAA 95153
Qy 3422 gaggcagcgtgaagcagatgagctgtgactagcaataatagggtttgtgtgtgtgtgt 3481
Db 95154 GAGGCACCGTGAAGACAGATGAGCTGTGGACTAGCAATATAGGGTTTGTGTGTGTGTGT 95213
Qy 3482 gggttgataaagcagtaatttggggtcattattgttccctgtgtgtgtgtgtgtgtgtgtgtgt 3541
Db 95214 GGTTTGATAAGCAGTATTTGGGGTCTATATTGTTTCTGTGTGTGTGTGTGTGTGTGTGTGT 95273
Qy 3542 actttgaagtatttatgttcttctcctcctcaattcattgtgtgtgtgtgtgtgtgtgtgtgt 3601
Db 95274 ACTTTGAAGTATTATATTGTTTCTTATCTCTCAATTCATATGTTGGTGTGATGAAATTCGAGGTT 95333
Qy 3602 gtcgtatttcttcagacttcgcacagacagattgtgtgataataataaattaggtaagataa 3661
Db 95334 GTCGTATATTCTTTCAGACTTCGCCACAGACAGATTGTGTGATAATAAATTTAGGTAAAGATAA 95393
Qy 3662 ttgttggggccatttttaggacaggtaaaataacatcagggttccagttcagttcgaattgc 3721
Db 95394 TTTGTTGGGCCATATTTTAGACAGGTAAATAACATCAGGTTCCAGTTCCTGTTGGAATTGC 95453
Qy 3722 aaggctaagaagtactgccccttttgt 3781

Db 95454 AAGGCTAAGAGTACTGCCCCCTTTTGTGTGTGTAGCAGTCAAAATCTATTATTCCACTGGCGC 95513
Qy 3782 atcatatgcagtgatatatgcctataataataagccaaggttcacaccattttgttttaga 3841
Db 95514 ATCATATGCAGTGATATATGCTTATAAATAAGCCATAGGTTACACACCATTTTGTTTTGA 95573
Qy 3842 caattgtcttttttcaagatgctttgtttcttcttccatatgaaaaaataatgcattttataa 3901
Db 95574 CAATTGTCTTTTCAAGATGCTTTGTTCTTTCATATAGAAAAAATGCATTTTATATA 95633
Qy 3902 ttcaagaagtcataagattttctgaagcgcacacgtgcatttttattttatgagctgtaagt 3961
Db 95634 TTCAGAAAGTCATAGATTCTCGAAGCGTCAACGTGTCATTTTATTATGAGCTGGTGAAGT 95693
Qy 3962 aactgtgtttactagcaggaatatttccaatttctacatttctactacatcttttcaaca 4021
Db 95694 AACTGTGGTTTACTAGCAGGAATATTTCCAATTTCTACCTTTTACTACATCTTTTCAACAA 95753
Qy 4022 gtaactttgtgaaatgagccagaagccctgagttggcagtggtggcccatagtg 4081
Db 95754 GTAACCTTTGTAGAAATGAGCCAGAAAGCCCAAGCCCTGAGTTGGCAGTGCCCACTAAGTGT 95813
Qy 4082 aaaaataaagtttccagaaacctt 4105
Db 95814 AAAATAAAAGTTTACAGAAACCTT 95837

RESULT 10

LOCUS G06463 2337 bp DNA linear STS 19-OCT-1995
DEFINITION human STS WI-7226, sequence tagged site.
ACCESSION G06463
VERSION G06463.1 GI:859708
KEYWORDS STS; STS sequence; primer; sequence tagged site.
SOURCE human STS derived from sequences in dbEST and the Unigene collection.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2337)
AUTHORS Hudson, T.
TITLE Whitehead Institute/MIT Center for Genome Research; Physically Mapped ESTs
JOURNAL Unpublished
COMMENT Contact: Thomas Hudson
Whitehead Institute/MIT Center for Genome Research
Whitehead Institute for Biomedical Research
9 Cambridge Center, Cambridge MA 02142 USA
Tel: 617 252 1900
Fax: 617 252 1902
Email: thudson@genome.wi.mit.edu
Primer A: GATCGAATTTTTCAGATGATTCG
Primer B: AAATGCCAGCAAAAGTCAC
STS size: 343
PCR Profile:
Presoak:
Denaturation:
Annealing: 56 degrees C
Polymerization:
PCR Cycles: 35
Thermal Cycler:
Protocol:
Template: 10 ng
Primer: each 5 pM
dNTPs: each 4 mM
Taq Polymerase: 0.025 units/ul
Total Vol: 20 ul
Buffer:
MgCl2: 1.5 mM
KCl: 50 mM


```

Db 1861 GACAGATTGCTGATAATAAATAGGTAAAGATAATTTGTTGGCCATATTTTAGGACAGGT 1920
QY 3689 aaataacatcaggttccagttgcttgaattgcaagtgctaaagtgactccctcttttg 3748
Db 1921 AAAATAACATCAGGTTCCAGTTCCTGCAATGCAAGCTAAGAAGTACTGCCCTTTTG 1980
QY 3749 tgttagcagtcgaatcattattccacgtgctgcacatcatatgcagtgatataatgcctataa 3808
Db 1981 TGTAGCAGTCAATCTATTATTCCACTGGCGCATCATATGCAGTGATATATGCTATAA 2040
QY 3809 tataaccataggttcacacccattgttttagacaattgctttttttaaagatgctttg 3868
Db 2041 TATAACCATAGGTTCCACACCATTTGTTTAGACAATTTGCTTTTTCGAAGATGCTTTG 2100
QY 3869 ttctttcatatgaaaaaagcattttataaattcagaagtcacagatttcttgaagcc 3928
Db 2101 TTTCTTTTCATATCAAAAAATGATTTTATAAATTCAGAAAGTCATAGATTTCTGAAGC 2160
QY 3929 gtcaacgtgcattttatttatgactggttaagtcaactggttttactagcaggaatattt 3988
Db 2161 GTCAACGTGCATTTTATTATGACTGGTAAGTAAGTAACTGTGTTTACTAGCAGGAATATT 2220
QY 3989 ccaattctacattactacatttttcaacagtaactttgtagaaatgagccagaagc 4048
Db 2221 CCAATTTCTACCTTTTACTACATCTTTTCAACAAGTAACTTTGAGAAATGAGCCAGAAGC 2280
QY 4049 caagccctgagttgcagtgcccaataagtgtaaaataaagtttacagaacacctt 4105
Db 2281 CAAGGCCCTGAGTTGCGACTGGCCCAATAAGTGTAATAAAGTTTACAGAACCTT 2337

RESULT 11
S63938
LOCUS S63938 1868 bp mRNA linear PRI 07-MAY-1993
DEFINITION A-type endothelin receptor [human, placenta, mRNA, 1868 nt].
ACCESSION S63938
VERSION S63938.1 GI:238635
KEYWORDS human placenta.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Adachi,M., Yang,Y.Y., Furulich,Y. and Miyamoto,C.
TITLE Cloning and characterization of cDNA encoding human A-type endothelin receptor
JOURNAL Biochem. Biophys. Res. Commun. 180 (3), 1265-1272 (1991)
MEDLINE 92062153
REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibsseq 63938] from the original journal article. This sequence comes from fig1.
FEATURES
    source
        1..1868
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
    gene
        1..1868
            /gene="A-type endothelin receptor, ETA"
    CDS
        74..1357
            /gene="A-type endothelin receptor, ETA"
            /note="This sequence comes from fig1; ETA"
            /codon_start=1
            /product="A-type endothelin receptor"
            /protein_id="AAB20278.1"
            /db_xref="GI:238636"
            /translation="MELICLRASFALVGVCSIDNPERYSTNLSNHVDDFTTFRGTE
            LSFVITHTQPNLVLPSPGSMHNYCPQOTKITSAPKYNIVTISCTIEFVGMVGNATLL
            RIYQNKMRPNALIASLALGLIYVIDLINVPKLLAGRWPFQHRDFVFLCKL
            PFLQKSVGIVTLNLCALSDVRVAVNSRVQIGIPLVTAIEIVSILTSILAI
            PEAGFVMPVEYGEQHKTMLANATSKMEFEYQDVQDKDWLFGFYFCMPLVCTAIFYT
            LMYCEMLNRRNGSIRIALSEHLKORREKAVTFCVLVIFALCWLPLSLRLKKTIVN
            EMDKNRCLESLFLIMDYIGINLATMNSCINPIALYFVSKFKNCFQSCGCCCCYQSK
            SLMTSVPMNGTSTIQWKHDQNNHNTDRSSHKDSMN"

```

BASE COUNT 512 a 419 c 389 g 548 t
ORIGIN

Query Match 44.3%; Score 1819.6; DB 9; Length 1868;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1844; Conservative 0; Mismatches 4; Indels 2; Gaps 2;

```

QY 421 aaaaagtgaaagtgtaaaaagcagcacaaagtgcaataagagatatatttctcctaaattgcct 480
Db 10 AAAAAGTGAAGGTGTAAGAGCAGCAAGTGCAATAAGAGATATTTCCTCAAAATTTGCCT 69
QY 481 caagatggaaaccccttggcctcagggcatccttttggctggcactgggttgatgtgtaaat 540
Db 70 CAAGATGGAACCCCTTTGCCCTCAGGGCATCCTTTTGGCTGGCACTGGTGGATGTGTAAT 129
QY 541 cagtataatacctgagagatacacagcacaaatctaaagcaatcatgtgatatgatttccaccac 600
Db 130 CAGTGTATAATCCTCGAGAGATACAGCACAAATCTAAGCAATCATCTGGATGATTATCCACCAC 189
QY 601 ttttcgtggcacagagctcagcttcctggttacactcacaacccaactaatttgggtcct 660
Db 190 TTTTCGTGGCACAGAGCTCAGCTTCCTCGGTACCACTCATCAACCCACTAATTTGGTCT 249
QY 661 acccagcaatggctcaatgcacaactattgcccacagcagactaaaattacttctcagcttt 720
Db 250 ACCCAGCAATGGCTCAATGCGACAACATATTGCCACAGCAGACTAAAATTACTTTCAGCTTT 309
QY 721 caaatataacactgtgatatcttctgtactattttctcgtgggaatgggtggggaatgc 780
Db 310 CAAATACATTAACACTGTGATATCTTGTACTATTATTTTCATCTGTGGGAATGTGGGGAATGC 369
QY 781 aactctgctcaggatcatttaccagacaataatgtatagggaatggcccaacgcgtgat 840
Db 370 AACTCTGCTCAGGATCATTTTACCAGAACAAATGTATAGGAATGGCCCCACGCGCTGAT 429
QY 841 agccagcttggccttggagacaccttatctatgtggtcattgactctcctcatcaatgatt 900
Db 430 AGCCAGCTCTGGCCTTTGGAGACCTTATCTATGTGGTCATTGATCTCCCTATCAATGATT 489
QY 901 taagctgctggctggcgctggccttttgatcacatgactttggcggtattttttgcaa 960
Db 490 TAAGCTGCTGGCTGGCGCTGGCGCTTTTGATCAACAATGACTTTTGGCGTATTCTTTGCCAA 549
QY 961 gctgttcccttttttcagaaagctcctgggtgggatacacogtctcaacotctgcgctct 1020
Db 550 GCTGTTCCCTTTTTCAGAAAGTCTCGGTGGGATCACCGTCTCAACCTCTGCGCTCT 609
QY 1021 tagttgtgacaggtacagagcagcttgcctcctcctgagtcglttcagggaattggattcc 1080
Db 610 TAGTGTGACAGGTACAGAGCAGTTCCTTCCCTGGAGTCTGGTTCAGGGAATTTGGGATTC 669
QY 1081 ttgtgtaactgccattgaaattgtctccatctggtactgcttcttcttctcctggccattcc 1140
Db 670 TTTGGTAACTGCCAATGAAATTTCTCCATCTGGATCTCTCTTTATCTCTGGCCATTC 729
QY 1141 tgaagcagattgctcgtctaatggtaacctttgaaataaggggtgacacagcataaaacctg 1200
Db 730 TGAAGCGATTGGCTTCGTCATGTACCCCTTTTGAATATAGGGGTGAACAGCATAAAAACCTG 789
QY 1201 tatgtctcaatgccacataaaaattcatgaggtctctaccaagatgtaaaggaactggtggct 1260
Db 790 TATGCTCAATGCCACATCAAAATTCATGAGTTCATCAAGATGTAAAGGACTGTGGTGGCT 849
QY 1261 cttcgggttctatttctgtatgccttggcttggctgcactgcgactcttctacacccctatgac 1320
Db 850 CTTTCGGGTCTATTCTGTATGCCCTTGGTGTGCACTGCGATCTTCTACACCCCTCATGAC 909
QY 1321 ttgtgagatgttgaacagagaaggaatggcagcttgagaattgcctcagtgaaacatttaa 1380
Db 910 TTGTGAGATGTTGAACAGAGAAGAA?GGCAGCTTGAGAATTGCCCTTCAGTGAACATCTTAA 969
QY 1381 gcagcgctgcagaagtgggcaaaaacagtttctctgcttgggtgtgtaatttttcttcttgcgtg 1440

```

[illegible]

RESULT 12

BTBTREC	BTBTREC	3216 bp	mRNA	linear	MAM 20-FEB-1991
LOCUS	Bovine mRNA for endothelin receptor.				
DEFINITION	X57765				
ACCESSION	X57765.1				
VERSION	GI:121				
KEYWORDS	endothelin receptor.				
SOURCE	cow.				
ORGANISM	Bos taurus				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;				
	Bovidae; Bovinae; Bos.				
REFERENCE	1 (bases 1 to 3216)				

Qy 542 agtgataatccctgagagatacagacaaatctaaagcaatcatgtgtgattgttaccact 601
Db 525 AGTGATAACCCCTGAAGACTACAGACACAAACCTTAACATCCATCGACAGTGTGCCACT 584
Qy 602 ttctgtgacagagactcagcttctcgtgtaccactcatcaaccacactaatgttgccta 661
Db 585 TTCCATGGGACAGAACTCAGCTTCGTGGTACCCACACATCAACCCACGAATTTGGCCCTC 644
Qy 662 cccagcaatggctcaatgcacaaactattgccacagcagactaaaattacttcagctttc 721
Db 645 CCCAGCAATGGCTCAATGCACAACTATTGCCACACAGACTAAGATTACTTCACCTTC 704
Qy 722 aaatacatcaactgtgatattctgtactatttctcgtggaatgtggggaatgca 781
Db 705 AAATACATTAACATGTGATATCATGTACTATTTTTCATCGTGGGAATGTGGGAATGCA 764
Qy 782 actctgctcaggaatcattaccagaacaaatgtatgaggaatggcccaacgcgtgata 841
Db 765 ACCCTGCTCAGGATCATTTACGAGAACAGTGCATGAGGAATGGCCCCAACGCGCTGATA 824
Qy 842 gccagttctgacctggagaccttatctatgtgttcattgtatctctcctatcaatgtatt 901
Db 825 GCCAGCCTTGCCCTGGGAGACCTTATCTATGTGTCAATTGATCTCCCTATCAATGTATT 884
Qy 902 aagctgtggtggcgctggcgtcttttgatcacaaatgactttggcgtatttctttgcaag 961
Db 885 AAGCTGTGGCTGGCGCTGGCCCTTTTGAACAAAATGACTTTGGCGTTTCTCTCTGCAAG 944
Qy 962 ctgttccccttttgcagaagctcctcgtgggagacccgtccctcaacctctgacctt 1021
Db 945 TTGTTCCCTTTTTCGAGAACTCTAGTGGGATCACCGTCCCTTAATCTCTGGCCCTTA 1004
Qy 1022 agtgtgacaggtacagagcagttgcctcctggagtcgtgttcagggaattgggattcct 1081
Db 1005 AGCCTTGACAGGTACAGAGCAGTTGCTCCTGGAGTCGTGTTTCAGGGAATTTGGGATCCC 1064
Qy 1082 ttggtaactgccattgaaattgtctccatctggtatctcgttcttcttctgaccttct 1141
Db 1065 TTGGTACAGCCCAATTGAGATTGTCTCCATCTGATCTGCTCTTTCTTTATCTCTGCCATCCCT 1124
Qy 1142 gaagcagattggctcgtcatggtaccctttgaatataggggtgaacagcataaaacctgt 1201
Db 1125 GAAGCCATCGCCTTTCATGAGTACCTTTTGAATACAGGGGTGCACAGCAGAACCTGT 1184
Qy 1202 atgtccaatgccacatcaaaatctaatgagttctaccagaagttaaaggactgtggctc 1261
Db 1185 ATGCTCAATGCCAGTCGAAATTCATGGAGTCTCTATCAGGATGTTAAGGACTGTGGCTG 1244
Qy 1262 ttccgggtctattctgtatgccttgggtgactgactgacttctcacaacctcactgact 1321
Db 1245 TTCGGGTCTACTCTCTGCATGCCCTGGTGTGCACTGCGATCTTTTACACCCCTCATGACT 1304
Qy 1322 tgtgagatgttgacagaagaatggcagcttgagaaattgccctcagtgaaactcttaag 1381
Db 1305 TGTGAGATGTTAAACGAGGAAGCGGAGCTGAGAAATTGCCCTCACTGAGCATCTGAAA 1364
Qy 1382 cagcgtgagaagtggcaaaaacagtttctgctgtgtgtgtaatttttgcctcttgcctgg 1441
Db 1365 CAGCGTCGAGAAGTGGCAAAAACAGTTTCTGCTGTGTTGTTAAATTTTGTCTGTGCTGG 1424
Qy 1442 ttccctctcaactaagccgtatattgaagaanaactgtgtatacgaanaatggacaagaac 1501
Db 1425 TTCCTCTCTCATTAAGCCGCATTTTGAAGAAAACCGCTGTATGATGAAATGGACACGAAC 1484
Qy 1502 cgaatgaaattacttacttctactgctcatgtgattacatcggtattacttggcaacc 1561
Db 1485 CGATGTGAATTAAGTTTCTTGTGCTCTCATGGATTACATCGGTATTAACCTTGGCAACC 1544
Qy 1562 atgaattcatgtataaaccccatagctctgtattttgtgagcaagaataatttaaaaattgt 1621
Db 1545 ATCAATTCATGTATCAACCAATAGCTCTATATTTTGTGAGCAAGAATAATTAATAATGT 1604
Qy 1622 ttccagtcactgctctgctgtgtgttaccagttccaaagctgatacctgggtcccc 1681

Db 1605 TTCCAGTCTTGCTCTGCTGCTGCTTTACCAGTCCAGAGCCTGATGACTTGGTGCCC 1664
Qy 1682 atgaacggaacaaagcatccagtggaagaaccacagatcaaaaacacacacacagaccgg 1741
Db 1665 ATGAATGGCACAGATCCAGTGGAGAACCAACCAACCAACCAACCAACCAAGAGG 1724
Qy 1742 agcagccataagagcagcatgaactgaccaccccttagaagcagcactcctcgttactccata 1801
Db 1725 AGCAGCCACAGACAGATATCAACT---AACTGAAGGAACAGTTATCAGTGTGCTTCAA 1781
Qy 1802 atcctctcgggaaaaaatacaagggaactgtgactcctcgggaactctctctctctgactcc 1861
Db 1782 GCCGACTGGAGGAAAAAATAACACAGCACTGTTTCCCACGAAATCTGTTC--CATCC 1838
Qy 1862 ttcttcttaattactcctccacacacagaagaatgctttccaaaacgcgaaggtagac 1921
Db 1839 TTTCTCAGACTCAGCCCCACCCCTGGAA--AAATGCTTTCCAAAACGTTGAAGGGTTG 1897
Qy 1922 tggtttatacccccaacacatctacgaatcgtactctttaaattgacttaatttaccatat 1981
Db 1898 ACTGCTTTATACCCACAAATTCACGAATCTTTACGCTCTTTGATTTATCTGATTTACCTAT 1957
Qy 1982 tctgctgttctgattcagcactaaataatggtgagctgg--gggagaatgaagactgt 2039
Db 1958 TCTGCTATGTTATTCAGCCCTTAAATAATGGTGGAGTTGGCAGGGGAAAGGAGATTGT 2017
Qy 2040 taaatgaacacagaagatatattactacttcttgcatgaaaaatagagc-tttcaagtacat 2098
Db 2018 TCAGTTAAACCCAGAGAGATATTTACCACTTTTCCGTTGAAATAGATTTTTCGAGTACAT 2077
Qy 2099 ggtcagctttctatgagcagttctggtgaatgttcaatgggaactggtccacatgaaacttt 2158
Db 2078 GGCRAACTTTCATGCTGGTCTGTAGAAATGTTCAATAAGACCAGTACCCTGAAAGATTT 2137
Qy 2159 agagattaacgacagaagatttctacttctttaaagtatttttgccttcagcgaacaa 2218
Db 2138 ATAGAGACTGTAAACAGATTTTCTTTTAACT--AAATTTGTTTTCGCCCCAACA 2196
Qy 2219 caatatgggtcaggtcacttttattgaaatgctcatttgggtgacagattttttaaactg 2278
Db 2197 CAGTATGAGCTGAAGTCTCCTTTTATTTGAAATATATTAATTCAGTSCCAGTATTTTAACTG 2256
Qy 2279 cataatgacctcaacatgattatttgaacttattacacatagtttgaaaaaagacaga 2338
Db 2257 CACAGTAGCCTACAGATGATTTGAGCTCATTTACACATATTTTTCGCCCCCAACAAAT 2316
Qy 2339 aaatagtagtattcaggtgagcaatagatttagtattttccacgctcactattttttta 2398
Db 2317 AAAACGGAAATTCAGGTGAGAAATAGTGTAGTGTTCATGCTCACTGGCTTATTTTFA 2376
Qy 2399 aacacacaaattcctaagctacacaaataactacagcccttaagcagcagctcgtat-gac 2457
Db 2377 AAACATAAATTCCTCAAGCTTCAGGAGCT--TACAGGAGCTTTAAAGCACAGACTGATATCC 2434
Qy 2458 acatttggcagtttaagtagatttactcaagaataatttttaagaactgtatttttttt 2517
Db 2435 ATATATGGCAGTTTAAATAGCTGCTATGCA---AACTTTTGGGAACCTATATTTCTTTT 2491
Qy 2518 taaatggtgttttattacaaagggacctgaacatggtttgtatgtttaattcaaaagtaa 2577
Db 2492 TTAATGGGTTCATATAGGAGATCTGTGCTTATTC---TGTAAATTTGAAAAATAA 2547
Qy 2578 tgcctcaatcagatagtttcttttcaagaagtcaaatcgttttttcaatgtaatttttga 2637
Db 2548 TGCTTGACTCAACAGTCTGTTTTCACAAAGCCTAGTTTTTAAATGGAAGCAATTTTATG 2607
Qy 2638 tgaataatcaatgccaagtacacaaatgttaattgtatgtcatttaactctgcctgaga 2697
Db 2608 TAAATCAAAATACCATTTACTAAAAATTTAAATGTTTGTATGCTTTAACTATGTCACAGA 2667
Qy 2698 ctttcagtcactgtatataagatctaaacacacacctcaagagaaaaagatacgaat-ttt 2756


```

Db 913 TCTTCGGGTTCTATTCTGTATGCCCTTGGTGTGCATCGGATCTTCTACACCTCATGA 972
Qy 1320 cttgtgagatgttgaacagaagaatggcagcttgagaattgcctcagtgaaacatctta 1379
Db 973 CTTGTGAGATGTTGACACAGAGGAATGGCAGCTTGAGATTGCCCTCAGTGAACATCTTA 1032
Qy 1380 agcagcgtcgaagaatggcaaaacagtttctgtggttgtaatttttgcctcttgc 1439
Db 1033 AGCAGCGTCGAGAAGTGGCAAAACAGTTTCTGCTGGTGTGAATTTTGTGCTCTTGGCT 1092
Qy 1440 gttccctctcacttaacgctatatattgaagaaactgtgtataacgaataaggacaaga 1499
Db 1093 GGTTCCTCTTCTATTAAAGCCGTATATTGAAGAAACTGTGTATAACGAGATGGACAAGA 1152
Qy 1500 accgatgtgaattacttagttcttactgtctcatggttacatcggtattaaacttgccaa 1559
Db 1153 ACCGATGTGAATTACTTACTTCTACTGCTCATGATTTACATCGGTATTAATCTGGCAA 1212
Qy 1560 caataatcatgtataaaaccccatagctctgtattttgtgagcaagaataattaaatt 1619
Db 1213 CCATGAATTCATGTATAACCCCATAGCTCTGTATTTTGTGAGCAAGAAATTTAAAAAT 1272
Qy 1620 gttccagtcagctctgtctgtctgtttaccagtcacaaagtctgtgacctcggtcc 1679
Db 1273 GTTCCAGTCATGCTCTGCTGCTGTGTACAGTCCAAAGTCTGTGACCTCGGTCC 1332
Qy 1680 ccatgaacggaacagatccagtgagagaaacacagatcaaaacacacacacagacc 1739
Db 1333 CCATGAACGGAACAGATCCAGTGGAGAACCCAGATCAAAACACCAACACACAGAGCC 1392
Qy 1740 ggaagcagccataagacagcatgaactgaccaccccttagaagcaactcctcggtactccca 1799
Db 1393 GGAGCAGCCATAAGGACAGCATGAAGTACCACCCCTTAGACGACTCTCTCGTACTCCCA 1452
Qy 1800 taactctctggagaaaaataacaaagcaactgactgactcgggaatctctctctgat 1859
Db 1453 TAATCTCTCGGAGAAAAAATFACAAGGCAACTGTGTGCTCGGGAATCTCTCTCTGAT 1512
Qy 1860 ccttctcttaattactccacaccccaagagaaatgctttccaaacccgaaggtag 1919
Db 1513 CTTCTCTCTTAATTCATCTCCACACCCNAGAAAGAAATGCTTCCAAACCCGACAGGTAG 1572
Qy 1920 actggtttatccaccacacacatctacgaatcgacttc-tttaattgatctaaattaca 1978
Db 1573 ACTGTTTATCCACCACCAACATCTACGAATCGTACTTCTTTTAAATGATCTAAATTACA 1632
Qy 1979 tatttcgctgtgttattcagcactaaaa 2007
Db 1633 TATTCTGCTGTGTATTTCAGCAGCTAAAA 1661

RESULT 14
S81539
LOCUS S81539 1359 bp mRNA linear PRI 03-AUG-1996
DEFINITION endothelin-A receptor-ET-AR [human, lung, mRNA Partial, 1359 nt].
ACCESSION S81539
VERSION S81539.1 GI:1478475
KEYWORDS
SOURCE human lung.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1359)
Miyamoto,Y., Yoshimasa,T., Arai,H., Takaya,K., Ogawa,Y., Itoh,H.
and Nakao,K.
Alternative RNA splicing of the human endothelin-A receptor
generates multiple transcripts
Biochem. J. 313 (Pt 3), 795-801 (1996)
96190719
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 176815] from the original journal article.
This sequence comes from Fig. 3A.

```

FEATURES

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

1. .1359

/gene="endothelin-A receptor"

1. .1284

/gene="endothelin-A receptor"

/note="ET-AR; This sequence comes from Fig. 3A"

/codon_start=1

/product="endothelin-A receptor"

/protein_id="AAB36325.1"

/db_xref="GI:1478476"

/translation="METICLRASFALVGCYSNDPERYSTNLSNHVDFTTFRGTE

LSFLVTHOPTNLVLPSPNGSMNYCPQOTKITSAFKYINTVICTLFIYGMVGNATLL

RIYXONKCMRNGPNALLASLALGDLIYVIDLPINVPFKLLAGRPDHDNDFGLCKL

FFFLQKSSGTVTLNLCUSDRVAVASRVSQVIGIPLVTAIEIVISWILSFILAI

PAIFQKVMVFFYEGEQHKTCLMLNATSKFEMVQYDKDWLFGYFCMPLVCKTAFT

LMTCEMLNRRNGSLRLALSEHLKORREAVTKFCLVIFALCWFLHLSILKTKTVN

EMDKNRCELLSFLLLMDYIGINLATMNSCINPIALYFVSKFKNCFQSLCCCCYOSK

SLMTSVPMNGTSTIQMKNHNDNNHTDRSSHKDSMN"

356 a 324 c 291 g 388 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity

Matches 1359; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

Qy 485 atggaaccccttgcctcagggcgcac

Db 1 ATGGAACCCCTTTGCTCAGGGCATC

Qy 545 gataatcctgagagatacagcacaatc

Db 61 GATAATCCTGAGAGATACAGCACAA

Qy 605 cgtggcacagagctcagctcctcgtt

Db 121 CGTGGCAGACAGCTCAGCTTCCTG

Qy 665 agcaatggctcaatgcacacactattg

Db 181 AGCAATGGCTCAATGCACAACTAT

Qy 725 tacattaaactgtgatatctgtactat

Db 241 TACATTAACACTGTGATATCTTGT

Qy 785 ctgctcagagcatattaccagaacaa

Db 301 CTGCTCAGGATCATTTACCAGAA

Qy 845 agtcttgccctggagacaccttatct

Db 361 AGTCTTGCCCTTGGAGACCTTAT

Qy 905 ctgctggctggcgctggcctttgat

Db 421 CTGCTGGCTGGCGCTGGCGCTTT

Qy 965 ttccctcttttcagaagtcctcgtgg

Db 481 TTCCCTCTTTTTCAGAAGTCCCT

Qy 1025 gttgacaggtacagagcagcttgc

Db 541 GTTGACAGGTACAGAGCAGTTC

Qy 1085 gtaactgccattgaaattgtctccat

Db 601 GTAACCTGCCATTGAAATGTCT

```
QY 1145 gcaatggcttcgctaccccttgaatataggggtgaacagacataaaacctgtatg 1204
DB 661 GCATTGGCTTCGCTATGCTAGCTTGAATATAGGGGTCAACAGCATAAACCTGTATG 720
QY 1205 ctcaatgccacatcaaaattcatggagttctaccagatgtaaaagactgggtgctcttc 1264
DB 721 CTCATGGCCACATCAAAATTCATGGAGTCTACCAAGATGTAAAGGACTGGTGGCTCTTC 780
QY 1265 gggcttatattgtatgccccctgggtgctgaactgcgattctctacacccctcatgacttgt 1324
DB 781 GGGCTTCTATTCTGTATGCCCTTGGTGTGCACCTGCGATCTTCTACACCCCTCATGACTGTG 840
QY 1325 gagattgaaagaagaaggaaatggcagacttgagaatggcctcagtaagaacatcttaagcag 1384
DB 841 GAGATGTGAACAGAGGAATGCGAGCTTGAGAAATGGCCCTCAGTGAACATCTTAAGCAG 900
QY 1385 cgtcgagaagtcgcaaaaacagctttctgctggtgtgtaattttctctcttctgctgcttc 1444
DB 901 CGTCGAGAAGTGCAAAACAGTTTTCTGCTGGTGTGAATTTTTGCTCTTTGCTGGTTC 960
QY 1445 cctctcacttaagccgcttatattgaagaaaactgtgtataacgaaatggagaaagaccga 1504
DB 961 CCTCTTCACTTAAGCCGTATATTGAAGAAACTGTGTATAACGAAATGGACAAGAACCGA 1020
QY 1505 tggcaattactagttcttactgcctcatgattacacatggtattacattggcacaacatg 1564
DB 1021 TGTGAATTACTTACTTACTTACTGCTCATGGATTACATCGGTATTAACTTGGCAACCATG 1080
QY 1565 aattcatgtataaaccccatagctctgattttgtgagcaagaatttaaaaattgtttc 1624
DB 1081 AATTCAATGATAAACCCCATAGCTCTGTATTTTGTGAGCAAGAAATTTAAATTTGTTTC 1140
QY 1625 cagtcacgtcctgctgctgctgttaccagtcacaaagtctgtatgacctggttcccoatg 1684
DB 1141 CAGTCATGCTCTGCTGCTGCTGTACCAAGAGTCTGTATGATGACCTCGTCCGCCCATG 1200
QY 1685 aacggaagaagcaccagtggaagaccagatcaaaaacacacacacagaccggagc 1744
DB 1201 AACGGAACAAGCATCCAGTGGAGAACCCAGATCAAAACACCAACACAGACCCGGAGC 1260
QY 1745 agcataagagacagcatgaactgaccaccccttagaagcactcctcggtactcccaatac 1804
DB 1261 AGCCATAAGACAGCATGACTGACCCACCTTTAGAGACATCTCTCGTACTCCCATATC 1320
QY 1805 ctctcgagagaaaaatcacaaggcaactgtgactccgg 1843
DB 1321 CTCTCGAGAGAAAAAATCACAAGGCAACTGTGACTCCGG 1359
```

```
RESULT 15
HUMEDNRA Homo sapiens endothelin receptor type A (EDNRA) mRNA linear PRI 31-DEC-1994
LOCUS Homo sapiens endothelin receptor type A (EDNRA) mRNA, complete cds.
DEFINITION L06622.1 GI:181956
ACCESSION L06622.1
KEYWORDS Homo sapiens (tissue library: endothelin receptor type A.
SOURCE Homo sapiens (tissue library: lambda ZAPII) heart, left ventricle
cdna to mRNA.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1310)
AUTHORS Elshourbagy,N.A., Korzan,D.R., Wu,H.L., Sylvester,D.R., Lee,J.A.,
Nuthalaganti,P., Bergsma,D.J., Kumar,C.S. and Nambi,P.
TITLE Molecular characterization and regulation of the human endothelin
receptors
JOURNAL J. Biol. Chem. 268 (6), 3873-3879 (1993)
MEDLINE 93179382
FEATURES Location/Qualifiers
source 1..1310
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_type="heart, left ventricle"
```

```

/tissue_lib="lambda ZAPII"
1..1310
/gene="EDNRA"
7..1290
/contig_start=1
/contig_end=1290
/product="endothelin receptor type A"
/protein_id="AAA58447.1"
/db_xref="GI:181957"
/translation="METLCRASFWLALVGVISDNPERYSNLSNVHVDFTTFRGTE
LSFLVTHQPTNLVPSGSMHNYCPOOTKTSFAKYINTVISCSTFIVGMVGNATLL
RIYONKCMRNPALIASLALGDLIYVVDLPINVKLLAGRWDFHNDGFLCKL
RFLQKSSVIGITVLNLCALSDVRYRAVASMVSVOGIGIPLVTAIGIASIWSLFIK
PRAIGFVMPFEYRGQHKTCMLNATSKFMESYODKDWLFGFYFCMLVCTAIFYT
LMTGMLNRRNGSLRIALSEHLKQRRVAKTVFCLVYIPALCWFLHLSRLIKKTVYN
ENDKRNKCELLSFLLLMDYIGINLATMNSCINPIALYFVSKFKFNCFQSCCLCCCTQSK
SLMTSVPMNGTSIQWNKHNDQNNHNDRSSHKDSMN"
BASE COUNT 337 a 312 c 284 g 377 t
ORIGIN
Query Match 31.6%; Score 1296.8; DB 9; Length 1310;
Best Local Similarity 99.5%; Pred. No. 4.3e-283;
Matches 1301; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 481 caagatggaaaacccttggcctcaggcgcaccccttttggctggcactgggttgatgtat 540
DB 3 CACCATGGAAACCCCTTGGCTCAGGCATCCTTTTGGCTGGCACCTGGTGGATGTGTAAT 62
QY 541 cagtataatcctcgagagatcacagcaaaatcctaaagcaatcatgttgatgtatccacc 600
DB 63 CACTGATATCTCGAGATAGACACAAATCTAAGCAATCATGTGGATGATTTACCAC 122
QY 601 tttctggcacagagctcagcttctcgtgtaccactcatcaaccactaatgtgctct 660
DB 123 TTTTCGTGGCACAGAGCTCAGCTTCTGTTTACCACTCATCAACCCACTAATTTGGTCT 182
QY 661 accagcaatggctcaatgcaaaactattgcccacagcagactaaaactactcagctt 720
DB 183 ACCCAGCAATGGCTCAATGCAACACTATTGCCACAGACACTAAAATTTACTTCAGTTT 242
QY 721 caatacattaaacactgtgatatctgtactatttttctcgtgggaatggtgggaatgc 780
DB 243 CAATACATTAACACTGTGATATCTTGTACTATTTCATCGTGGGAATGCTGGGAATGC 302
QY 781 aactctgctcagatcatattaccagaaacaaatgtatgaggaaatggccccacgcgtcat 840
DB 303 AACTCTGCTCAGGATCATTTTACCAGAAATGTATGAGGAATGGCCCCAACGCGCTGAT 362
QY 841 agccagcttgccttggagacacttatctatgtggtcattgactcctcctcactaatgtatt 900
DB 363 AGCCAGCTCTTGGCCCTTGGAGACCTTATCTATGTGGTCAATTGATCTCCCTATCAATGATT 422
QY 901 taagctgctgctggcgctggccttttggatcacaaatgactttggcgctattcttctgcaa 960
DB 423 TAAGCTGCTGCTGGCGCTGGCCCTTTTGATCACAATGACTTTGGCGTATTCTTTCGAA 482
QY 961 gctgttccccctttttggagaagtcctcgtggggatcacccgctcactcactcgcgtct 1020
DB 483 GCTGTTCCCTTTTGGCAGAAAGTCTCGTGGGATCACCCTCTCAACCTCTCGCTCT 542
QY 1021 tagtggacaggtacagacagcttgcctcctggagtcggttcagggaattgggattcc 1080
DB 543 TAGTGTGTGACAGGTACAGACAGCTTGGCTCTCTGAGTGTGCTTTCAGGAAATGGGATTC 602
QY 1081 ttggtaactgcatgaaattgtctcactgtgactcctcttcttcttctcctggcattcc 1140
DB 603 TTTGGTAACCTGCCATTAATAATGCTCCATCTGGATCCTGTCTCTTTATCTGCCATTC 662
QY 1141 tgaagcagattggcttcgtcatggtaccctttgaatatagggtgaaacagacataaaacctg 1200
DB 663 TGAAGCATTTGGCTTCGTCATGTTACCTTTGAATATAGGGGTGACAGCATAAACCTG 722
```


OM of: US-09-931-157-1 to: Issued_Patents_AA.* out_format : pfs
Date: Oct 1, 2002 1:06 AM
About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framed_n2p.model -DEV=xlp
-Q/cgn2_1/USPTO.spool/US09931157/runat_30092002_110420_11267/app_query.fasta_1.8657
-DB=Issued_Patents_AA -OFMT=fastan -SUFFIX=rai -GAPOPT=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPT=0.000
-GAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -XGAPOP=10.000 -XGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=blotsum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USRP=US09931157 -CGN1_1.85 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-931-157-1
Query length: 4105
Database: Issued_Patents_AA.*
Database sequences: 231628
Database length: 24425594
Search time (sec): 89.890000

score_list:

Sequence	Strd Orig	ZScore	EScore Len	Documentation
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-121-446-2		2282.00	4003.35	3.9e-216 427
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-118-270-40		1571.00	2748.78	3.8e-146 333
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US93-08528-40		1571.00	2748.78	3.8e-146 333
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-121-446-4		1240.00	2159.87	1.8e-113 442
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-117-361C-1		1225.00	2133.07	5.5e-112 452
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-118-270-41		899.00	1559.56	6.3e-80 350
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US93-08528-41		899.00	1559.56	6.3e-80 350
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-07-937-609-17		465.00	790.78	3.8e-37 384
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-029-170-17		465.00	790.78	3.8e-37 384
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US92-02091-2		465.00	790.78	3.8e-37 384
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US92-02091-4		455.50	773.98	3.3e-36 384
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-07-937-609-18		454.50	772.02	4.1e-36 390
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-029-170-18		454.50	772.02	4.1e-36 390
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US92-02091-6		454.50	772.02	4.1e-36 390
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US92-02091-8		434.50	736.65	3.9e-34 390
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-910-092-2		414.50	701.72	3.5e-32 376
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-279-590A-2		411.50	696.41	7.0e-32 376
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-120-772-2		410.00	693.03	1.0e-31 399
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-118-270-42		404.50	685.69	3.2e-31 328
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US93-08528-42		404.50	685.69	3.2e-31 328
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-118-270-43		404.50	685.08	3.3e-31 345
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US93-08528-43		404.50	685.08	3.3e-31 345
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-118-270-37		398.50	675.73	1.2e-30 311
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US93-08528-37		398.50	675.73	1.2e-30 311
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-465-687A-2		391.50	655.08	8.6e-30 613
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-030-970-2		391.50	655.08	8.6e-30 613
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US94-11843-2		391.50	655.08	8.6e-30 613
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-959-381A-2		379.50	635.36	1.2e-28 542
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-959-381A-1		367.00	614.70	1.9e-27 481
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-430-286A-2		324.00	542.31	2.8e-23 356
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-149-093A-5		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-911-245-5		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-889-108-2		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-120-701B-2		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-387-707-16		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-510-473-5		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:PCT-US94-10358-2		323.50	540.07	3.4e-23 398
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-454-549-3		322.50	538.51	4.2e-23 391
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-454-552-3		322.50	538.51	4.2e-23 391
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-188-275A-3		322.50	538.51	4.2e-23 398

/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-351-198-3 + 322.50 538.30 4.2e-23 39
/cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-09-113-426-3 + 322.50 538.30 4.2e-23 39
/cgn2_6/ptodata/2/1aa/6A_COMB.pap:US-08-676-351-4 + 316.50 527.90 1.6e-22 33
/cgn2_6/ptodata/2/1aa/6A_COMB.pap:US-08-889-108-8 + 315.50 525.85 2.1e-22 40
/cgn2_6/ptodata/2/1aa/6A_COMB.pap:US-08-188-275A-2 + 315.50 525.85 2.1e-22 4
seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pap:US-08-121-446-2
seq_documentation_block:
; Sequence 2, Application US/08121446
; Patent No. 6313276
; GENERAL INFORMATION:
; APPLICANT: IMURA, HIROO
; APPLICANT: NAKAO, KAZUWA
; APPLICANT: NAKANISHI, SHIGETADA
; TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121,446
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,684
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 29900-20324.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 427 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-121-446-2
alignment_scores:
Quality: 2282.00 Length: 427
Ratio: 5.344 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 100.000
alignment_block:
US-09-931-157-1 x US-08-121-446-2
Align seg 1/1 to: US-08-121-446-2 from: 1 to: 427
485 ATGGAAACCTTTGGCTCAGGCGATCCTTTGGCTGGCCTGGTGGATG 534
1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyCy 17
535 TGTAACTAGTGAATACTCTGAGAGATACAGCAAAATCAAGCAATCATG 584
17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisv 34
585 TGGATGATTTTACCATTCTTCTGGGACAGAGCTCAGCTTCCTGGTTACC 634
|||||

34 aAspAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
635 ACTCATCAACCCACTAATTTGGTCTCACCAGCAATGGCTCAATGCACAA 684
51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMetHisAs 67
685 CTATTGCCACAGACAGACTAAATTTACTTACGTTTCAATACATACAA 734
67 nTyCysProGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
735 CTGTGATATCTTGACTATTTTCATCGTGGGAATGCTGGGAATCAACT 784
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
785 CTGCTCAGGATCATTTACCAGAACAAATGTATAGGAATGGCCCCAAGC 834
101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
835 GCTGATAGCCAGTCTTGCCCTTGAGACCTTATCTATGTGTGTCATGTC 884
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
885 TCCCTATCAATGTATTTAAGCTGCTGGCTGGCGCTGGCCTTTGATCAC 934
134 euProIleAsnValPheLysLeuLeuAlaGlyArgTyrProPheAspHis 150
935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTTTTTCGCAAGTC 984
151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSe 167
985 CTCGGTGGGATACCGTCTCAACCTCTGGCTCTTGTAGTGTGACAGGT 1034
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
1035 ACAGACGAGTTGCTCCTGGAGTCTGTGTACAGGAATGGCATTCCTTGG 1084
184 yArgAlaValAlaSerTyrPheArgValGlnGlyIleGlyIleProLeu 200
1085 GTAACGCAATGAAATGTCTCCATCTGGATCTGCTCTTTATCCTGGC 1134
201 ValThrAlaIleGluIleValSerIleTyrPheLeuSerPheIleLeuAl 217
1135 CATTCCTGAACGATTTGGCTTCGTCATGGTACCTTTGAATATAGGGGTG 1184
217 aileProGluAlaIleGlyPheValMetValProPheGluTyrArgGlyG 234
1185 AACAGCATAAACCTGTATGCTCAATGCCACATCAAAATTCATGAGTTC 1234
234 luGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
1235 TACAAGATGTAAAGGACTGGTGGCTCTTCGGGTCTTATTTCTGTATGCC 1284
251 TyrGlnAspValLysAspTyrTyrPheGlyPheTyrPheCysMetPr 267
1285 CTGGTGTGACTCGCATCTTCTACACCTCATGACCTGTGTGAGATGTGA 1334
267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
1335 ACAGAAGGAATGGCAGCTTGAGATTCGCCCTCAGTGAACATCTTAAGCAG 1384
284 snArgArgsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
1385 CGTCGAGAAGTGGCAAAACAGTTTCTGCTTGTGTTGTAATTTTGTCT 1434
301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
1435 TTGCTGTTCCCTTTCACCTTAAGCCGTATATTCGAAGAAACACTGTGTATA 1484
317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 ACGAATGGACAAACCGATGTGAATTACTTAGTTTCTACTGCTCATG 1534
334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350

1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAACCCCAT 1584
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
1585 ACCTCTGTATTTTGTGAGCAAGAATTTAAAAATTTTCCAGTCATGCC 1634
367 eaIaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TCTGTGCTGCTGTTACCAGTCCAAAGTCTGATGACCTCGGTCCCATG 1684
384 euCysCysCysTyrGlnSerLysSerLeuMetThrSerValPromet 400
1685 AACGGAACAAGCATCCAGTGGAGAACCCAGATCAAAACAAACCAACAC 1734
401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnHisAsnTh 417
1735 AGACGGAGCAGCCATAGGACAGCATGAAC 1765
417 rAspArgSerSerHisLysAspSerMetAsn 427
seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-118-270-40

seq_documentation_block:
; Sequence 40, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIVAK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-118-270-40

alignment_scores:
Quality: 1571.00 Length: 347
Ratio: 4.864 Gaps: 5
Percent Similarity: 93.084 Percent Identity: 85.303


```

789 TCAGGATCATTTACAGACAAATATGAGGAATGGCCCCAACCGCTG 838
|||||
123 euArgileIleYrLysAsnLysCysMetArgAsnGlyProAsnIleLeu 139
839 ATAGCCAGCTTCGCCCTTGGAGACCTTATCTATGTGGTCATTGATCTCCC 888
|||||
140 IleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleValIleAspIlePr 156
889 TATCAATGTATTAAAGCTGCTGGCTGGCGCTGGCTTTTGATCACAATG 938
|||||
156 IleAsnValYrLysLeuLeuAlaGluAspPrPro..... 168
939 ACTTTGGCGTATTTCTTTCAGAGCTGTTCCTCTTTTTCAGAGCTCGG 988
|||||
169 ..PheCylAlaGluMetCysLysLeuValProPheIleGlnLysAlaSer 184
989 GTGGGATCACCCTCTCAACCTCTGCCCTCTAGTGTGACAGTACAG 1038
|||||
185 ValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrAr 201
1039 AGCAGTTGCTCTCTGAGTCGTGTCAGGGAATGGGATTCCTTGGTAA 1088
|||||
201 GalValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpT 218
1089 CTGCCAFTGAATGTCTCCATCTCGATCCTCTCTCTTATCTGCGCCATT 1138
|||||
218 hrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVal 234
1139 CTTGAAGCATTTGGCTTCTGCTCATGCTACCTTTGAATATAGGGTCAACA 1188
|||||
235 ProGluAlaIleGlyPheAspIleIleThrMetAspTyrLysGlySerTy 251
1189 GCATAAAACCTCTATGCTCAATGCCACATCAAAA.....TTCATGGAGT 1232
|||||
251 IleuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetGlnP 268
1233 TCTACCAAGATTAAGGACTGGTGGCTCTTCGGGTCTTATTTCTGTATG 1282
|||||
268 heTyrLysThrAlaLysAspTrpLeuPheSerPheYrPheCysLeu 284
1283 CCTTGGTGTGCTGCGATCTTCTACACCCCTCATGCTTGAGATGTT 1332
|||||
285 ProLeuAlaIleThrAlaPheYrThrLeuMetThrCysGluMetLe 301
1333 GAACAGAAGATGGCAGCTTCAGAAATGCCCTCAGTGAACATCTTAAGC 1382
|||||
301 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 317
1383 AGCGTCGAGAAGTGGCAAAACAGTTTCTGCTTGTGTTGTAATTTTGTCT 1432
|||||
317 InArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 333
1433 CTTTGTGCTGCTCTCTTACCTTAAGCCCTATATTGAAGAAACCTGTGTA 1482
|||||
334 LeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTy 350
1483 TAACGAAATGGACAGACCGATGTGAATTACTTACTTCTTCTTACTGCTCA 1532
|||||
350 rAsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuValL 367
1533 TGGATTACATCGGTATTACTTGGCAACCATGAATTCATGTATAACCCC 1582
|||||
367 euAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 383
1583 ATAGCTCTGTATTTCTGAGCAAGAAATTTAAATTTGTTTCCAGTCATG 1632
|||||
384 IleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCy 400
1633 CTTCTGCTGCTGTGTACAGTCCCAAAAGTCTGTGACCTCGGTCGCCCA 1682
|||||
400 sleuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGluG 416

```

```

1683 TCACGGGAACAGCATCCAGTCGAGAGACCAGCATCAAAACACACACAC 1732
|||||
416 LuLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyrAsp 431
1733 ACAGACCGGAGCAGCCATAAGACACAGC 1759
|||||
432 AsnPheArgSerSerAsnLysTyrSer 440
seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-117-361C-1

seq_documentation_block:
: Sequence 1, Application US/08117361C
: Patent No. 5773223
: GENERAL INFORMATION:
: APPLICANT: Venkatakrishna, Shyamala
: APPLICANT: Tekamp-Olson, Patricia
: TITLE OF INVENTION: Endothelin B1 (ETB1) Receptor Polypeptide
: TITLE OF INVENTION: Compositions, Methods, and Uses Thereof
: NUMBER OF SEQUENCES: 27
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Chiron Corporation
: STREET: 4560 Horton Street
: CITY: Emeryville
: STATE: California
: COUNTRY: USA
: ZIP: 946087
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25 &
: WORDWARE: Word 5.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/117,361C
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Chung, Ling-fong
: REGISTRATION NUMBER: 36,482
: REFERENCE/DOCKET NUMBER: 0945.001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (510) 923-2704
: TELEFAX: (510) 655-3542
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 452 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: -
: LOCATION: 199..208
: OTHER INFORMATION: /note= "Decapeptide Insert"
US-08-117-361C-1

alignment_scores:
Quality: 1225.00 Length: 369
Ratio: 3.939 Gaps: 6
Percent Similarity: 84.282 Percent Identity: 61.789

alignment_block:
US-09-931-157-1 x US-08-117-361C-1 ..
Align seg 1/1 to: US-08-117-361C-1 from: 1 to: 452

689 TCACCACAGCAGCACTAAAATTTACTTCAGCTTTCAATATCAATCACTGT 738
|||||
90 CysGlnGlyProIleGluIleLysGluThrPheLysTyrIleAsnThrVa 106
739 GATATCTTGTACTATTTTCATCGTGGGAATGGTGGGAATGCAACTCTGC 788
|||||
106 lValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeuL 123

```



```
911 .....GCTGGCGCTGGCCTTTTGCATCAC 934
59 hrGlyThrIleValSerGlnHisAlaAlaLeuPheGluTrpHisAla 75
935 AATGACTTTGGCGTATTTCTTGAAGCTGTCCCTTTTTCGACGAAGTC 984
76 ValAspProGlyCysArgLeuValPheIleGlnLysAl 92
985 CTCGGTGGGATCACCGCTCTCAACCTCTGGCCTCTTGTAGTTGACAGGT 1034
92 aserValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgT 109
1035 ACAGAGCAGTTGCCCTCGGAGTCGTCTCAGGGAATTGGGATTCCTTG 1084
109 yrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLys 125
1085 GTAACCTGCATTGAATGTCTCCATCTGGATCGCTCTTATCTGCGC 1134
126 TrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuAl 142
1135 CATTCCTGAAGCGATTGGCTTGCATGTTACCCCTTTGAATATAGGGGTG 1184
142 aValProGluAlaIleGlyPheAspThrThrSer...AspTyrLysGlyL 158
1185 AACACGATAAACCTGTATGCTCAATGCCACATCAAA.....TTCATG 1228
158 ysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMet 174
1229 GAGTTCTACCAAGATGTAAGACATGGTGGCTCTTCGGGTTCTATTCTG 1278
175 PheTyrLysThrAlaAlaLysAspTrpPheAlaPheTyrPheCys 191
1279 TATGCCCTTGGTGCATGCGATCTTCTACACCTCATGACTTGTGAGA 1328
191 sLeuProLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGluM 208
1329 TGTTCACAGAGGAATGGCAGCTTGAGAAATGCCCTCAGTCAACATCTT 1378
208 etLeuArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeu 223
1379 AAGCAGCGTCGAGAAGTGCAGAAACAGTTTCTGCTTGGTGTAAATTT 1428
224 LysGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPh 240
1429 TGCTCTTGTGCTGGTCCCTCTTCACTTAAGCCGTATATTGAAGAAACTG 1478
240 eaLeuLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrL 257
1479 TGTATAACGAATGGACAGAACCGATGTGAATTACTTACTTCTTACTG 1528
257 eutyAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeu 273
1529 CTCATGATTACATCGGTATTAACTGGCAACCATCAATTCATGTATAAA 1578
274 ValLeuAspTyrIleGlyIleAsnMetAlaSerIleAsnSerCysIleAs 290
1579 CCCCATGACTCTGTATTTTGTGAGCAGAAATTTAAATTTGTTTCCAGT 1628
290 nProIleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysS 307
1629 CATGCTCTGCTGCTGCTGTTTACCAGTCCAAAAGTCTGATGACCTCGTC 1678
307 erCysLeuCysCysTrpCys...GlnThrPheGluGluLysGlnSerLeu 322
1679 CCCATCAACGGAACAGCATCCAGTGAAGAACACCGATCAAAACAAACCA 1728
323 GluGluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTy 338
1729 CAACACAGACCGGAGCAGCATTAAGGACAGC 1759
338 rAspAsnPheArgSerSerAsnLysTyrSer 348
```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pcp: PCT-US93-08528-41

```
seq_documentation_block:
; Sequence 41, Application PC/TUS9308528
; GENERAL INFORMATION:
; APPLICANT: New York University
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEWMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/08528
; FILING DATE: 09-SEP-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2 PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; INFORMATION FOR SEQ ID NO: 41:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 350 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; PCT-US93-08528-41
```

alignment_scores:
Quality: 899.00 Length: 311
Ratio: 3.596 Gaps: 7
Percent Similarity: 80.386 Percent Identity: 57.235

alignment_block:
US-09-931-157-1 x PCT-US93-08528-41 ..

Align seg 1/1 to: PCT-US93-08528-41 from: 1 to: 350
865 TATCATGTGGTCATTGATCTCCCTATCAATGATTTAAAGCTGCTG... 910
43 PheLeuCysGly-LeuValLeuThrLeuAspPheLeuGlyLeuLeuVal 59
911GCTGGCGCTGGCCTTTTGCATCAC 934
59 hrGlyThrIleValSerGlnHisAlaAlaLeuPheGluTrpHisAla 75
935 AATGACTTTGGCGTATTTCTTGAAGCTGTCCCTTTTTCGACGAAGTC 984
76 ValAspProGlyCysArgLeuValPheIleGlnLysAl 92
985 CTCGGTGGGATCACCGCTCTCAACCTCTGGCCTCTTGTAGTTGACAGGT 1034
92 aserValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgT 109
1035 ACAGAGCAGTTGCCCTCGGAGTCGTCTCAGGGAATTGGGATTCCTTG 1084
109 yrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLys 125

```
1085 GTAACGCCATTGAATGTCTCCATCTGTCATCTGTCCTCTTTATCTGSC 1134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
126 TrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuAl 142
1135 CATTCTCTGAAGCCATGGCTTCCTCATGTCATGGTACCTTTGATATAGGGGTG 1184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
142 avaiProGluAlaIleGlyPheAspThrThrSer...AspTyrLysGlyL 158
1185 AACAGCATAAACCTCTATGCTCAATGCCACATCAAAA.....TTCATG 1228
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
158 ysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMet 174
1229 GAGTTCTACAGATGTAAGGACTGGTGGCTCTTCGGGTCTATTCTG 1278
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
175 PheTyrLysThrAlaAlaLysAspTrpTrpLeuPheAlaPheTyrPheCy 191
1279 TATGCCCTTGGTGTGCACCTCGCATCTTACACCCCTCATGCTGGAGA 1328
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
191 sleuProLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGluM 208
1329 TGTGTAACAGAAGGATGGCAGCTTGAGAAATGCCCTCAGTGAACATCTT 1378
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 etLeuArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeu 223
1379 AAGCAGCTGCGAAGTGGCAAAACAGTTTCTGCTTGGTTCTTAATTT 1428
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
224 LysGlnArgGluValAlaLysThrValPheCysLeuValLeuValPh 240
1429 TGCCTCTTCTGCTGCTCCCTTCACCTTACCGCTATATGAGAAACACTG 1478
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 exAlaLeuCystrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrL 257
1479 TGTATAACGAAATGGACAAGACCGATGTGAATTACTTACTTACTTACTG 1528
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
257 euTyrAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeu 273
1529 CTCATGGATTACATCGGTATTACTTGGCAACCATGAATTCATGTATAAA 1578
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
274 ValLeuAspTyrIleGlyIleAsnMetAlaSerIleAsnSerCysIleAs 290
1579 CCCCATAGCTCTGTTTGTGCGCAAGAAATTTAAATTTGTTTCCAGT 1628
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 nProIleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLys 307
1629 CATGCTCTGCTGCTGCTGTTACCAAGTCCAAAGTCTGATGACCTCGGTC 1678
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
307 erCysLeuCysCysTrpCys...GlnThrPheGluGluLysGlnSerLeu 322
1679 CCCATGAACGGAACAGCATCCAGTGGAGAACCCAGCATCAAAACACCA 1728
||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
323 GluGluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTy 338
1729 CAACACAGACCGGACGCCCATTAAGACAGC 1759
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
338 rAspAsnPheArgSerSerAsnLysTyrSer 348
seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-07-937-609-17
```

```
seq_documentation_block:
; Sequence 17, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
```

```
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/937,609
FILING DATE: 19920902
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/831,248
FILING DATE: 07-FEB-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/861,769
FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/928,033
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: mouse gastrin-releasing peptide receptor
US-07-937-609-17

alignment_scores:
Quality: 465.00 Length: 342
Ratio: 2.076 Gaps: 7
Percent Similarity: 65.497 Percent Identity: 32.456

alignment_block:
US-09-931-157-1 x US-07-937-609-17 ..
Align seg 1/1 to: US-07-937-609-17 from: 1 to: 384
719 TTCAATACATTAAACACTGTGATATCTGTACTATTTTTCATCTGTGGAAT 768
||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
40 PheIleTyrValIleProAlaValTyrGlyLeuIleIleValIleGlyLe 56
769 GGTGGGAATGCAACTGCTGCTCAGGATCATTTACCAGAACAAATGTATGA 818
:|||||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
56 uIleGlyAsnIleThrIleIleLysIlePheCysThrValLysSerMeta 73
819 GGAATGGCCCAACCGCTGATAGCAGCTGTGCCCTTGAGACCTTATC 868
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
73 rGsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLeu 89
869 TATGTGTTCATTCCTCCCTATCAATGATTTTAAGCTGCTGGCTGGCG 918
:||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
90 LeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspAr 106
919 CTGGCCCTTTTGATCACAATGACATTTGGCGTATTTCTTTGCAAGCTGTCC 968
||||| |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
106 gtrpLeu.....PheGlyArgIleGlyCysLysLeuIleP 118
969 CCTTTTTCGAGAAGTCTCGGTGGGATCACCGCTCTCAACCTCTCGCT 1018
|||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
118 ropheIleGlnLeuThrSerValGlyValSerValPheThrLeuThrAla 134
1019 CTTAGTGTTCACAGGTACAGACAGTTGCCCTCTGGAGTCGTGTCAGGG 1068
|||||: |||: |||: |||: |||: |||: |||: |||: |||: |||: |||:
```

```

COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/029,170
  FILING DATE: 19930310
  CLASSIFICATION: 435
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/937,609
  FILING DATE: 02-SEP-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/928,033
  FILING DATE: 11-AUG-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/861,769
  FILING DATE: 01-APR-1992
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 07/831,248
  FILING DATE: 07-FEB-1992
ATTORNEY/AGENT INFORMATION:
  NAME: BENT, Stephen A.
  REGISTRATION NUMBER: 29,768
  REFERENCE/DOCKET NUMBER: 40399/166 NIHD
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (703)836-9300
  TELEFAX: (703)683-4109
  TELEX: 899149
  INFORMATION FOR SEQ ID NO: 17:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 384 amino acids
      TYPE: AMINO ACID
      TOPOLOGY: linear
IMMEDIATE SOURCE:
  CLONE: mouse gastrin-releasing peptide receptor
US-08-029-170-17

alignment_scores:
  Quality: 465.00      Length: 342
  Ratio: 2.076        Gaps: 7
  Percent Similarity: 65.497      Percent Identity: 32.456

alignment_block:
US-09-931-157-1 x US-08-029-170-17  ..

Align seg 1/1 to: US-08-029-170-17 from: 1 to: 384

719 TTCAATATACATTAACACACTGTGTATATCTTGCTACTATTTTCATCTCGGGAAT 768
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40 PheleIyVaiIlePrOnAlaValIyrgLyLeuIleIleValIleGlyLe 56
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
56 uIleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 73
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
769 GGTGGGGAATCAACTGCTGCTCAGGATCATTTTACCAGAACAAATGTATGA 818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
819 GGAATGGCCCCAACCGCGTGATAGCCAGTCTTTCGCCCTTGAGACCTTATC 868
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
73 rGAsnValPrOAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeu 89
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
869 TATGTGGTCATTGATCTCCCTATCAATGATTTTAAGCTGCTGCTGGCGG 918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
90 LeuLeuValThrCyAlaPrOValAspAlaSerLysrLysrLeuAlaAsp 106
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
919 CTGGCGCTTTTGATCAACAATGACTTTTGGCGCTATTTCTTTGCAAGCTGTTC 968
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 gTrpLeu.....PheGlyArgIleGlyCysLysLeuIleP 118
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
969 CCTTTTTCGAGAAGTCCTCGGCGGGGATCACCGTCTCTCAACCTCTGCGCT 1018
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```
118 roPheileGlnLeuThrSerValGlyValSerValPheThrLeuThraLa 134
1019 CTTAGTCTTACAGGTACAGAGAGTGTCCCTCGGAGTCTGTTCAGGG 1068
135 LeuSerAlaAspArgTyrLysAlaIleValArgProMetAspIleGlnAl 151
1069 AATTGGGATTCCTTGGTAAGTCTCCATCTCCATCTGGATCC 1118
151 aSerHisAlaLeuMetLysIleCysLeuLysAlaAlaLeuIlePrlEv 168
1119 TGTCTTTATCCCTGGCATTCTTGAAGCGATTGGCTTCGTTCATGGTACCC 1168
168 alSerMetLeuLeuAlaIleProGluAlaValPheSerAspLeuHisPro 184
1169 TTTGAATATAGGGGTGAACAGCATAAAACCTGTATGCTCAATGCCACATC 1218
185 PheHisValLysAspThrAsnGlnThrPheIleSerCysAlaProTyrPr 201
1219 AAAATTCATGGACTTCTACCAAGATGTAAGGACTGGTGGCTCTTCGGGT 1268
201 oHisSerAsnGluLeuHisProLysIleHisSerMetAlaSerPheLeuV 218
1269 TCTATTCTGTATGCTTGGTGTGCTGCGACTCTTCTACACCCCTCATG 1318
218 alPheTyrValIleProLeuAlaIleSerValTyrTyrPheIle 234
1319 ACTGTGTGATGTGTGAACAGAAGG.....AATGGCAGCTT 1353
235 AlaArgAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnII 251
1354 GAGATTCGCCCTCAGTGAACATCTTAAGCAGCTCGAGAGTGGCAAAA 1403
251 ehIsVal.....LysLysGlnIleGluSerArgLysArgLeuAlaLysT 266
1404 CAGTTTCTGCTGGTGTAAATTTTGGCTTCTTGGTGGTCCCTCTTCAC 1453
266 hrValLeuValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHis 282
1454 TTAAGCCGTATATTGAAGAAACTGTGTATAAGAAATGGACAGAACC 1503
283 ValIleTyrLeuTyrArgSerTyrHisTyrSerGluValAspThrSer.. 298
1504 ATGTGAATTACTAGTTCTTCTTACTGCTCATGATTACATCGGTATTAACT 1553
299 .....MetLeuHisPhe.....ValThrSerIleCysAlaHisLeuL 311
1554 TGGCAACCATGATATATAAACCCTAGCTCTGTATTGTGTGAGC 1603
311 euAlaPheThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSer 327
1604 AAGAAATTTAAATTTGTTCCAGTCATGCCCTGCTGCTGTGTGTATACCA 1653
328 LysSerPheArgLysGlnPheAsnThrGlnLeuLeuCysCys..... 341
1654 GTCCAAAGTCTGATGACCTCGCTGCCATGAACGGAACAACATCCAGT 1703
342 -GlnProGlyLeuMetAsn.ArgSerHisSerThrGlyArgSerThrThr 357
1704 GGAAGAACCCAGCATCAAAACAACC 1727
358 CysMetThrSerPheLysSerThr 365
seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:pct-US92-02091-2
seq_documentation_block:
; Sequence 2, Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Battey Jr., James F.
; APPLICANT: Corjay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwin P. Ching
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02091
; FILING DATE: 19920313
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/426,150
; FILING DATE: 24-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,659
; FILING DATE: 05-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34090
; REFERENCE/DOCKET NUMBER: A-0092C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-266-7476
; TELEFAX: 415-266-7400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02091-2

alignment_scores:
  Quality: 465.00      Length: 342
  Ratio: 2.076        Gaps: 7
Percent Similarity: 65.497 Percent Identity: 32.456

alignment_block:
US-09-931-157-1 x PCT-US92-02091-2
Align seg 1/1 to: PCT-US92-02091-2 from: 1 to: 384
719 TTCAATACATTAAACACTGTGATATCTTGTACTATTCTTCATCTCGTGGAA 768
40 PheIleTyrValIleProAlaValTyrGlyLeuIleIleValIleGlyLe 56
769 GGTGGGAATGCAACTCTGCTCAGGATCATTTACCAGAACAAATGTATGA 818
56 uileGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 73
819 GGAATGGCCCCAACGCGCTCATAGCCAGTCTGCCCTTGAGACCTTATC 868
73 rgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLe 89
869 TATGTGCTCATTTGATCTCCCTATCAATGTATTAAAGCTGCTGGCTGGCG 918
90 LeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspAr 106
919 CTGGCCCTTTTGTATCAACATGACTTTGGCGGTATTCTTTTGCAGCTGTTC 968
106 gTrpLeu.....PheGlyArgIleGlyCysLysLeuIleP 118
969 CCTTTTTCAGAGAGTCTCGGTGGGATCATCCGCTCCTCAACCTCTCGGCT 1018
118 roPheIleGlnLeuThrSerValGlyValSerValPheThrLeuThraLa 134
1019 CTTAGTCTTACAGGTACAGAGAGTGTCCCTCGGAGTCTGTTCAGGG 1068
```



```
1075 GATTCCTTTGGTAACCGCCATTGCAATTTGCTCCATCTGATCCTGTCTCT 1124
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
152 sAlaLeuMetLysIleCysLeuLeuLysAlaAlaPheIleTrpIleLeuSerM 169
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1125 TTATCTGCGCCATCTCTGAAGCATTGGCTCTGCTCATGTTGACCTTTT 1171
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
169 eLeuLeuAlaIleProGluAlaValPheSerAspLeuHisProPheHis 185
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1172 ...GAATATAGGGGTGAACAGCATAAACTGTATGCTCAATGCCACATC 1218
    ||| : : : : : : : : : : : : : : : : : : : : : : : :
186 GluGluSerThrAsnGlnThrPheIleSerCysAlaProTyrProHisSe 202
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1219 ARAATTCATGGAGTCTACCAAGATTAAGGACTGCTGCTCTTCGGGT 1268
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
202 rAsn.....GluLeuHisProLysIleHisSerMetAlaSerPheLeuV 217
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1269 TCTATTCTGTATGCTGCTGGTGTGCACTGCGATCTTACACCCCTCATG 1318
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
217 alpheTyrrValIleProLeuSerIleIleSerValTyrTyrPheIle 233
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1319 ACTTGTGAGATGTTGAACAGAAGG.....AATGGCAGCTT 1353
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 AlAlLysAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnI 250
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1354 GAGATTCGCCCTCAGTGAACATCTTAAGCAGCGTCGAGAGTGGCAAAA 1403
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
250 eHisVal.....LysLysGlnIleGluSerArgLysArgLeuAlaLysT 265
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1404 CAGTTTCTGCTGTGTGTAATTTTGTCTTTGCTGTTGCTCTCTTCTCAC 1453
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
265 hrValLeuValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHis 281
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1454 TTAAGCCGTATATTGAAGAAAACCTGTATACGAATGCAAGAAGACCG 1503
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
282 ValIleTyrLeuTyrArgSerTyrHisTyrSerGluValAspThrSerMe 298
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1504 ATGTCGAATTACTTACTTCTTCTACTGCTCATGATGATACATCGTATACT 1553
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
298 tLeuHisPheValThrSerIleCysAlaArg.....LeuL 310
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1554 TGGCAACCATGATGATGATGATGATGATGATGATGATGATGATGATG 1603
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
310 euAlaPheThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSer 326
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1604 AAGAAATTTAAATTTGTTCCAGTCATGCTGCTGCTGCTGCTGCTGCTG 1645
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
327 LysSerPheArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGl 343
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1646 .....TGTACAGTCCCAAAAGCTGTGATGACCT 1673
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
343 yLeuIleIleArgSerHisSerThrGlyArgSerThrThrCysMetThrs 360
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1674 CGGTCCCATGACGAGCAAGCATC 1699
    ||||| : : : : : : : : : : : : : : : : : : : : : : : :
360 eLeuLysSerThrAsnProSerVal 368
```

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-07-937-609-18

```
seq_documentation_block:
; Sequence 18, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/937,609
; FILING DATE: 19920902
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat neuromedin B receptor
; US-07-937-609-18
```

alignment_scores:

Quality:	454.50	Length:	408
Ratio:	1.959	Gaps:	12
Percent Similarity:	56.863	Percent Identity:	30.392

alignment_block:

US-09-931-157-1 x US-07-937-609-18 ..

Align seg 1/1 to: US-07-937-609-18 from: 1 to: 390

```
650 AATTTGGTCTACCCAGCAATGGCTCA..... 676
      ||||| ||||| : : : : : : : :
      8 AsnLeuSerLeuProThrGluAlaSerGluSerGluLeuGluProGluVa 24

677 ....ATGCACAACTATTGCCACAGCAGACTAAATTTACTTCAGCTTCA 722
      : : : : : : : : : : : : : :
      24 lTrpGluAsnAspPheLeuProAspSerAspGlyThrThrAla..... 38

723 AATACATTAACTGTGATATCTTGTACT..... 751
      : : : ||||| ||||| : : :
      39 .....GluLeuValIleArgCysValIleProSerLeuTyrLeuIle 52

752 ATTTTCATCGTGGGAATGGTGGGAATGCAACTCTGCTCAGGATCATTTA 801
      ||||| : : : : : : : : : :
      53 lIleIleSerValGlyLeuLeuGlyAsnIleMetLeuValLysIlePheLe 69

802 CCAGAACAAATCTATGAGGAATGGCCCAACCGCTGATAGCAGTCTTG 851
      ||||| : : : : : : : : : :
      69 uThrAsnSerThrMetArgSerValProAsnIlePheIleSerAsnLeuA 86

852 CCCTTGGAGACCTTATCTATGTGGTCATTGATCTCCCTATCAATGTTT 901
      || ||||| : : : : : : : :
      86 laAlaGlyAspLeuLeuLeuLeuThrCysValProValAsp..... 100

902 AAGCTGCTGGCGCTGGCGCTTTTGTATCACAATAGCTTTGGCGTATT 951
      ||||| : : : : : : : :
      101 .....AlaSerArgTyrPhePheAspGluTrpValPheGlyLysLe 114
```

```
952 TCTTTGCAAGCTGTCTCCCTTTTTCGACGAAGTCCCTGGTGGGATCACCG 1001
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
114 uclyCysLysLeuIleProAlaIleGlnLeuThrSerValGlyValServ 131
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1002 TCCTCAACCTCTGCCTCTTACTGTGTACAGGTACAGACGACTTGCCTCC 1051
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
131 alpheThrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsn 147
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1052 TGGAGTCGTGTTGACGGAATTCGGATTCCTTTGGTAACCTGCCATTGAAT 1101
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
148 PrometAspMetGlnThrSerGlyValValLeuThrPheSerLeuLysAl 164
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1102 TGTCTCCATCTGCATCCTCTCTATCTCTCTGCGCATTCCTGAAGCGATT. 1150
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
164 aValGlyIleTrpValValSerValLeuLeuAlaValProGluAlaValP 181
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1151 .....GGCTTC...GTC 1159
181 heSerGluValAlaArgIleGlySerSerAspAsnSerSerPheThrAla 197
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1160 ATGGTACCCCTTTGAATATAGGGGTGAACAGCATAAACCTGTATGCTCAA 1209
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
198 CysIleProTyrProGlnThrAspGluLeuHis..... 208
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1210 TGCCACATCAAAATTCATGGAGTCTTACCAAGATCTAAAGGACTGTGGC 1259
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
209 .....ProLysIleHisSerValLeuI 216
1260 TCTTCGGGTCTATTCTGTATGCCCTTGCTGTGCACCTCGCATCTCTAC 1309
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
216 lePheLeuValTyrPheLeuIleProLeuValIleSerIleTyr 232
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1310 ACCCTCATCAGCTGTGAGATGTGAACAGAGAAGTGGCAGCTTGAGAT 1359
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
233 ...TyrHisIleAlaLysThrLeuIleArgSerAlaHisAsnLeuProG 248
1360 TGCCCTCAGTGAACATCTTAACAG.....CCTCGAAGAAGTGG 1397
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
248 yGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLysArgLeuA 265
1398 CAAAAACATTTCTGCTGGTGTGTAATTTTCTCTCTTGTGGTTCCT 1447
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
265 lalysIleValLeuValPheValGlyCysPheValPheCysTrpPhePro 281
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1448 CTCACCTTAAGCGGTATATTGAAGAAACTGTGTATACGAAATGGACAA 1497
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
282 AsnHisIleLeuTyrLeuTyrArgSerPheAsnTyrLysLysIleAspr 298
1498 GAACCGATGTGAATTACTTAGTTTCTTACTGCTCATGGATTACATCGTA 1547
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
298 oSerLeuGlyHisMetIleValThrLeu.....ValAlaA 310
1548 TTAACCTTGCACCAACCATGATTCATGATATAACCCCATAGCTCTGTTT 1597
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
310 rgValLeuSerPheSerAsnSerCysValAsnProPheAlaLeuTyrLeu 326
1598 GTGACGAAGAAATTTAAAATTTCTTCCAGTCATGCTCTGCTGC..... 1642
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
327 LeuSerGluSerPheArgLysHisPheAsnSerGlnLeuCysCysGlyG 343
1643 TGCTCTTACCAGTCCAAAGTCTG..... 1666
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
343 nLysSerTyrProGluArgSerThrSerTyrLeuLeuSerSerSerAlav 360
1667 .....ATGACCTCGTCCCATCAACGGAACAGCATCCAGTGAAGAAC 1711
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 alArgMetThrSerLeuLysSerAsnAlaLysAsnValValThrAsnSer 376
1712 CAGCATCAAAACAACACACACACA 1735
    : |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
377 valLeuLeuAsnGlyHisSerThr 384
```

```
seq_name: /cgn2_6/ptodata/2/iaa/6B_COMB.pep:US-08-029-170-18
seq_documentation_block:
: Sequence 18, Application US/08029170
: Patent No. 6169173
: GENERAL INFORMATION:
: APPLICANT: WANK, Stephen A.
: TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
: TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
: NUMBER OF SEQUENCES: 32
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Foley & Lardner
: STREET: 1800 Diagonal Road, Suite 500
: CITY: Alexandria
: STATE: VA
: COUNTRY: USA
: ZIP: 22313-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/029,170
: FILING DATE: 19930310
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/937,609
: FILING DATE: 02-SEP-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/928,033
: FILING DATE: 11-AUG-1992
: APPLICATION NUMBER: US 07/861,769
: FILING DATE: 01-APR-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/831,248
: FILING DATE: 07-FEB-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: BENT, Stephen A.
: REGISTRATION NUMBER: 29,768
: REFERENCE/DOCKET NUMBER: 40399/166 NIHD
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)836-9300
: TELEX: (703)683-4109
: INFORMATION FOR SEQ ID NO: 18:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 390 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: IMMEDIATE SOURCE:
: CLONE: rat neuromedin B receptor
: US-08-029-170-18
alignment_scores:
Quality: 454.50 Length: 408
Ratio: 1.959 Gaps: 12
Percent Similarity: 56.863 Percent Identity: 30.392
alignment_block:
US-09-931-157-1 x US-08-029-170-18 ..
Align seg 1/1 to: US-08-029-170-18 from: 1 to: 390
650 AATTTGGTCTCTACCCAGCAATGGCTCA..... 676
8 AsnLeuSerLeuProThrGluAlaSerGluSerGluLeuGluProGluVa 24
677 ....ATGCACAACATTTGCCACAGCAGACTAAATTTACTTTCAGTTTCA 722
24 lTrpGluAsnAspPheLeuProAspSeraspGlyThrAla..... 38
```



```
723 AATACATTAAACAGTGTGATCTCTGTACT..... 751
39 .....GluLeuValIleArgCysValIleProSerLeuTyrLeuIle 52
752 ATTTTCATCGTGGGAATGCTGGGAGTGAACCTCTCAGGATCATTTA 801
53 IleIleSerValGlyLeuLeuGlyAsnIleMetLeuValIlePheLe 69
802 CCAGAACAAATGTATGAGGAATGGCCCAACGCGCTGTAGTACGAGCTTG 851
69 uthrAsnSerThrMetArgSerValProAsnIlePheIleSerAsnLeuA 86
852 CCCTGGGACACTTATCTATGTGGCTCATGATCTCCTATCAATGATTT 901
86 laalaglyAspLeuLeuLeuLeuThrCysValProValAsp..... 100
902 AAGCTGCTGGCTGGCGCTGGCTTTTGATCAACAATGACTTTGGCGTATT 951
101 .....AlaSerArgTyrPhePheAspGluTrpValPheGlyLysLe 114
952 TCTTTCAGAGCTGTTCCTCTTTTTCAGAGTCTCTCGTGGGATCACCG 1001
114 uclyCysLysLeuIleProAlaIleGlnLeuThrSerValGlyValSerV 131
1002 TCTCAACCTCTGCGCTCTTAGTGTGACAGGTACAGACAGTGTGCTCC 1051
131 alPheThrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsn 147
1052 TGGAGTCTGTTTCAGGAATTCGGATCTCTTGGTAACGTCCATTCGAAT 1101
148 ProMetAspMetGlnThrSerGlyValValLeuTrpThrSerLeuLysAl 164
1102 TGTCTCCATCTGATCTGCTCTTATCTCTGCGCATCTCTGAAGCGATT. 1150
164 aValGlyIleTrpValValSerValLeuLeuAlaValProGluAlaValP 181
1151 .....GGCTTC...GTC 1159
181 heSerGluValAlaAlaArgIleGlySerSerAspAsnSerSerPheThrAla 197
1160 ATGGTACCCTTTGATATAGGGGTGAACAGCATAAACCTGTATGCTCAA 1209
198 CysIleProTyrProGlnThrAspGluLeuHis..... 208
1210 TGCCACATCAAAATTCATGGAGTTCTACCAAGATGTAAAGGACTGTGTGC 1259
209 .....ProLysIleHisSerValLeuI 216
1260 TCTTCGGGTTCTATTTCGTATGTCCTTGGTGTGCACCTCGCATCTTCTAC 1309
216 lePheLeuValTyrPheLeuIleProLeuValIleIleSerIleTyrTyr 232
1310 ACCCTCATGACTGTGAGATGTTGACAGAGGAGGATGGCAGCTTGAGAT 1359
233 ...TyrHisIleAlaLysThrLeuIleArgSerAlaHisAsnLeuProGl 248
1360 TGCCCTCAGTGAACATCTTAAGCAG.....CGTCGAGAGTGG 1397
248 yGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLysArgLeuA 265
1398 CAAAAACAGTTTCTGCTGGTGGTGAATTTTTCCTCTTGTGCTGCTCCCT 1447
265 laLysIleValLeuValPheValGlyCysPheValPheCysTrpPhePro 281
1448 CTTACATTAAACCGGTATATTGAGAAACTGTGTATACGAATATGACAA 1497
282 AsnHisIleLeuTyrLeuTyrArgSerPheAsnTyrLysGluIleAspPr 298
1498 GAACCGATGTGAATTACTTACTTTCTTACTGCTCATGATTACATCGGTA 1547
298 oSerLeuGlyHisMetIleValThrLeu.....ValAlaIa 310
```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US92-02091-6

seq_documentation_block:
; Sequence 6, Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Battey Jr., James F.
; APPLICANT: Corjay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwin P. Ching
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02091
; FILING DATE: 19920313
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/426,150
; FILING DATE: 24-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,659
; FILING DATE: 05-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34090
; REFERENCE/DOCKET NUMBER: A-0092C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-266-7476
; TELEFAX: 415-266-7400
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US92-02091-6

alignment_scores:
Quality: 454.50 Length: 408
Ratio: 1.959 Gaps: 12
Percent Similarity: 56.863 Percent Identity: 30.392

alignment_block:

US-09-931-157-1 x PCT-US92-02091-6

Align seg 1/1 to: PCT-US92-02091-6 from: 1 to: 390

```
650 AATTGGTCTACCCAGCAATGGCTCA..... 676
||||| ||||| ||||| ||||| ||||| |||||
8 AsnLeuSerLeuProThrGluAlaSerGluLeuGluProGluVa 24
677 ....ATGCACAACTATTCCCCACACGACACTAAATATTACTTCAGCTTCA 722
||||| ||||| ||||| ||||| ||||| |||||
24 lTrpGluAsnAspPheLeuProAspSerAspGlyThrThrAla..... 38
723 AATCATTTAACACTGTGATATCTTGTACT..... 751
||| ||||| ||||| ||||| ||||| |||||
39 .....GluLeuValIleArgCysValIleProSerLeuTyrLeuIle 52
752 ATTTTCATCGGGGAAATGGTGGGAATGCAACTCTGCTCAGATCATTTA 801
||||| ||||| ||||| ||||| ||||| |||||
53 IleIleSerValGlyLeuLeuGlyAsnIleMetLeuValLysIlePheLe 69
802 CCAGACAAATGCTATGAGGATGGCCCAACGCGCTGATAGCAGCTCTTG 851
||||| ||||| ||||| ||||| ||||| |||||
69 uThrAsnSerThrMetArgSerValProAsnIlePheIleSerAsnLeuA 86
852 CCTTTGGAGACCTTATCTATGTGGTCAATTGATCTCCCTATCAATGATTT 901
|| ||||| ||||| ||||| ||||| ||||| |||||
86 IaAlaGlyAspLeuLeuLeuLeuLeuThrCysValProValAsp..... 100
902 AAGCTGCTGGTGGCGCTGGCCTTTTGTATCACAATGACTTTGGCGTATT 951
||||| ||||| ||||| ||||| ||||| |||||
101 .....AlaSerArgTyrPhePheAspGluTrpValPheGlyLysLe 114
952 TCTTTGCAAGCTGTTCCTCTTTCCAGAACTCCCTCGGTGGGATCAGCG 1001
||||| ||||| ||||| ||||| ||||| |||||
114 uClyCysLysLeuIleProAlaIleGlnLeuThrSerValGlyValSerV 131
1002 TCCTCAACTCTGCGCTCTTGTGTGACAGTACAGAGCAGTTCCTCC 1051
||||| ||||| ||||| ||||| ||||| |||||
131 alPheThrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsn 147
1052 TGGACTCTGTTCAGGAATGGGATTCCTTGTGTAAGTGCATGAAAT 1101
||||| ||||| ||||| ||||| ||||| |||||
148 PrometAspMetGlnThrSerGlyValValLeuTrpThrSerLeuLysAl 164
1102 TGTCTCCATCTGGATCCTGCTCTTATCTGCGCATTCCTGAAGCGATT. 1150
||||| ||||| ||||| ||||| ||||| |||||
164 aValGlyIleTrpValValSerValLeuLeuAlaValProGluAlaValP 181
1151 .....GGCTTC...GTC 1159
181 heSerGluValAlaArgIleGlySerSerAspAsnSerSerPheThrAla 197
1160 ATGTGTACCTTTGAATATAGGGTCAACAGCATAAACCTGTATCCTCAA 1209
||||| ||||| ||||| ||||| ||||| |||||
198 CysIleProTyrProGlnThrAspGluLeuHis..... 208
1210 TCCACATCAAAATTCATGGAGTTCTACCAAGATGTAAGGACTGGTGGC 1259
209 .....ProlLysIleHisSerValLeuI 216
1260 TCTTCGGGTCTTATTCTGTATGCTTGGTGTGCACTGCGATCTTCTAC 1309
||||| ||||| ||||| ||||| ||||| |||||
216 lePheLeuValTyrPheLeuIleProLeuValIleIleSerIleTyrTyr 232
1310 ACCCTCAGCTGTGACAGATTTGACACAGGAATGGCAGCTTGAGAAAT 1359
||||| ||||| ||||| ||||| ||||| |||||
233 ...TyrHisIleAlaLysThrLeuIleArgSerAlaHisAsnLeuProGl 248
1360 TCCCTCAGTGAACATCTTAACAG.....CCTCAGGAAGTGG 1397
||||| ||||| ||||| ||||| ||||| |||||
248 yGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLysArgLeuA 265
```

```
1398 CAAAACAGTTTCTGCTTGGTTGTAATTTTGTCTTCTTGTGTTCCCT 1447
||||| ||||| ||||| ||||| ||||| |||||
265 IaLysIleValLeuValPheValGlyCysPheValPheCysTrpPhePro 281
1448 CTTCACTTAAGCCGTATATTGAAGAAACTGTGTATTAACGAANTGCACA 1497
||||| ||||| ||||| ||||| ||||| |||||
282 AsnHisIleLeuTyrLeuTyrArgSerPheAsnTyrLysGluIleAspPr 298
1498 GAACGATGTGAATTTACTTAGTTCTTCTGCTCATGATTACATCGGTA 1547
||| ||||| ||||| ||||| ||||| |||||
298 oSerLeuGlyHisMetIleValThrLeu.....ValAlaA 310
1548 TTAACCTGGCAACCATGAATTCATGATAAACCCCATAGCTCTGTATTT 1597
||||| ||||| ||||| ||||| ||||| |||||
310 rgValLeuSerPheSerAsnSerCysValAsnProPheAlaLeuTyrLeu 326
1598 GTGACCAAGAAATTTAAAATTTTCCAGTCATGCCCTCTGCTGC..... 1642
||||| ||||| ||||| ||||| ||||| |||||
327 LeuSerGluSerPheArgLysHisPheAsnSerGlnLeuCysCysGlygl 343
1643 .TGCTGTTACCAGTCCAAAAGTCTG..... 1666
||| ||||| ||||| ||||| ||||| |||||
343 nLysSerTyrProGluArgSerThrSerTyrLeuLeuSerSerSerAlav 360
1667 .....ATGACCTGGTCCCATGAACGGAACCAAGCATCCAGTGGGAAC 1711
||||| ||||| ||||| ||||| ||||| |||||
360 alArgMetThrSerLeuLysSerAsnAlaLysAsnValValThrAsnSer 376
1712 CACGATCAAAACACCAACACACA 1735
||||| ||||| ||||| ||||| ||||| |||||
377 ValLeuLeuAsnGlyHisSerThr 384
seq_name: /cgn2_5/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US92-02091-8
seq_documentation_block:
; Sequence 8' Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Battey Jr., James F.
; APPLICANT: Corjay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESS: Edwin P. Ching
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02091
; FILING DATE: 19920313
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/426,150
; FILING DATE: 24-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,659
; FILING DATE: 05-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34090
; REFERENCE/DOCKET NUMBER: A-0092C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-266-7476
; TELEFAX: 415-266-7400
```

[illegible]

OM of: US-09-931-157-1 to: A_Geneseq_032802.* out_format : pfs

Date: Oct 1, 2002 1:04 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+n2p.model -DEV=xlp
-Q/cgn2_1/USPTO_spool/US09931157/runat_30092002_110420_11256/app_query.fasta_1.8657
-DB=A_Geneseq_032802 -QFMT=fastan -SUFFIX=tag -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DEL0P=6.000 -DELEXT=7.000 -START=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09931157-ecnl_1_398 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-931-157-1

Query length: 4105

Database: A_Geneseq_032802.*

Database sequences: 747574

Database length: 111073796

Search time (sec): 320.970000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1993.DAT:AA30885	2282.00	4059.77	1.3e-218	42	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA3078	2282.00	4059.77	1.3e-218	42	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AB56351	2277.00	4050.81	4.1e-218	42	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA48719	1571.00	2787.54	1.2e-147	33	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA02691	1371.00	2787.54	1.2e-147	33	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA25394	1240.00	2190.51	1.6e-114	44	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1993.DAT:AA30886	1240.00	2190.51	1.6e-114	44	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AB56352	1235.00	2181.54	5.2e-114	44	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1998.DAT:AA59436	1225.00	2163.34	5.3e-113	45	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA48720	899.00	1581.70	1.7e-80	33	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA02692	899.00	1581.70	1.7e-80	33	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1991.DAT:AA12125	465.00	802.20	4.1e-37	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1991.DAT:AA12399	465.00	802.20	4.1e-37	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA27507	465.00	802.20	4.1e-37	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA86620	465.00	802.20	4.1e-37	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA27508	456.50	786.95	2.9e-36	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA86621	454.50	783.18	4.6e-36	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA27509	451.50	777.80	9.1e-36	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AB56359	450.50	776.19	1.1e-35	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA27510	434.50	747.31	4.6e-34	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA57956	434.50	747.31	4.6e-34	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AB56369	429.50	738.34	1.4e-33	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA89359	415.50	707.84	4.6e-32	61	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA89359	414.50	711.88	4.4e-32	37	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA90667	414.50	706.05	5.8e-32	61	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA07002	411.00	705.41	1.0e-31	38	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA58919	410.00	703.10	1.3e-31	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA87512	410.00	703.10	1.3e-31	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1992.DAT:AA27511	406.00	695.92	3.2e-31	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA48721	404.50	695.57	4.1e-31	32	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA02693	404.50	695.57	4.1e-31	32	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA48722	404.50	694.97	4.2e-31	34	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA02694	404.50	694.97	4.2e-31	34	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AB56333	404.00	692.34	5.1e-31	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2001.DAT:AA07638	404.00	692.34	5.1e-31	39	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA48716	398.50	685.44	1.6e-30	31	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA02688	398.50	685.44	1.6e-30	31	
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1996.DAT:AA898261	394.50	670.18	5.7e-30	61	

/SIDSL/gcgdata/geneseq/geneseq-embl/AA1998.DAT:AA37799 + 394.50 670.18 5.7e-3
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1998.DAT:AA55030 + 379.50 644.74 1.7e-3
/SIDSL/gcgdata/geneseq/geneseq-embl/AA2000.DAT:AA90665 + 368.00 625.54 2.2e-3
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1998.DAT:AA55029 + 367.00 623.75 2.8e-3
/SIDSL/gcgdata/geneseq/geneseq-embl/AA1999.DAT:AA32886 + 367.00 623.75 2.8e-3

seq_name: /SIDSL/gcgdata/geneseq/geneseq-embl/AA1993.DAT:AA30885

seq_documentation_block:

ID AAR30885 standard; Protein; 427 AA.

AC AAR30885;

XX 11-MAY-1993 (first entry)

DT Eta receptor.

DE Human; ETa; ETb; endothelin; receptor; transmembrane domain; N tail;

KW extracellular; cytoplasmic; C tail; post translational; bovine;

KW modification; ET-1 receptor; antagonist; circulatory system.

XX Homo sapiens.

OS

Key

Peptide

Protein

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Domain

Location/Qualifiers
1..20 "Signal peptide"
21..427
81..105 "Mature peptide"
117..139 "Transmembrane domain I"
160..181 "Transmembrane domain II"
204..229 "Transmembrane domain III"
257..278 "Transmembrane domain IV"
308..332 "Transmembrane domain V"
347..372 "Transmembrane domain VI"
379 "Transmembrane domain VII"
29 "N-glycosylated site"
62 "N-glycosylated site"
379 "N-glycosylated site"
383 "May be palmitoylated as in beta-2-adrenergic receptor"
385..388 "One base may be palmitoylated as in beta-2-adrenergic receptor"
289 "May be phosphorylated with serine/threonine kinases"
391 "May be phosphorylated with serine/threonine kinases"
421 "May be phosphorylated with serine/threonine kinases"
92EP-0306347.
91JP-0172828.

PA (SHIO) SHIONOGI SEIYAKU KK.
 XX
 PI Imura H, Nakanishi S, Nakao K;
 XX
 DR WPI: 1993-010677/02.
 DR N-PSDB; AAQ34583...
 XX
 PT Human Eta and Etb endothelin receptors - for measuring endothelin
 PT and screening for endothelin antagonists
 XX
 PS Claim 1; Fig 1; 39pp; English.
 XX
 CC The sequences given in AAR30885-86 represent the human Efa and Etb
 CC endothelin receptors respectively. Efa is a 427 amino acid protein
 CC with a molecular weight of 48,726. Etb comprises 442 amino acids and
 CC has a molecular weight of 49,629. Efa has a higher affinity for
 CC endothelin (ET)-1 and ET-2, whereas Etb has no selectivity for ET-1,
 CC ET-2 or ET-3. The receptors each contain seven transmembrane domains
 CC and have an extracellular N tail and a cytoplasmic C tail. There are
 CC several potential sites for post translational modification, these
 CC sites are identical to those of bovine ET-1 receptor. Efa cDNA is
 CC 91.2% homologous to bovine ET-1 receptor cDNA and Etb cDNA is 61.8
 CC homologous to that of bovine Efa receptor. The receptor proteins are
 CC useful as reagents for measuring the amount of Et or screening for
 CC antagonists of the Et receptor when studying the circulatory system.
 XX
 SQ Sequence 427 AA;

alignment_scores:
 Quality: 2282.00 Length: 427
 Ratio: 5.344 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-931-157-1 x AAR30885

Align seg 1/1 to: AAR30885 from: 1 to: 427

485 ATGGAACCCCTTTGGCTCAGGSCATCCTTTGGCTGGCAGCTGGTGGATG 534
 |||||
 1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyC 17
 |||||
 535 TGTATCATGATATCTCTGAGAGATACAGCACAAATCTAAGCAATCATG 584
 |||||
 17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisV 34
 |||||
 585 TGGATGATTTACACACTTTTCGTGGCCAGAGCTCAGCTTCCTGTTACC 634
 |||||
 34 AlaAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
 |||||
 635 ACTCATCAACCCACTAATTTGGTCTCTACCCAGCATGGCTCAATGCACAA 584
 |||||
 51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMetHisAs 67
 |||||
 685 CTATTGCCACAGCAGACTAAATTTACTTTCAGCTTTCAATATACATTAACA 734
 |||||
 67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84
 |||||
 735 CTGTGATATCTGTACTATTTTCATCGTGGGAATGGTGGGAATGCAACT 784
 |||||
 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
 |||||
 785 CTGCTCAGGATCATTTACCAGAACAAATGTATGAGGAATGSCCCCAAGCC 834
 |||||
 101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
 |||||
 835 GCTGATAGCCAGTCTTGCCTCTGGAGACTTATCTATGCTGCTCATGATC 884
 |||||
 117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
 |||||
 885 TCCCTATCAATGATTTAAGCTGCTGGCTGGCGCTGGCTTTTGTATCAC 934
 |||||

134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHis 150
 |||||
 935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTTTTTCAGAAAGTC 984
 |||||
 151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysse 167
 |||||
 985 CTGGTGGGGATCACCCTCTCAACCTCTCGGCTCTTAGTGTTCACAGGT 1034
 |||||
 167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArg 184
 |||||
 1035 ACAGACAGTTGGCTCTCTGGAGTCGTCTCAGGGAATTTGGGATTCCTTTG 1084
 |||||
 184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProleu 200
 |||||
 1085 GTAACGTGCATTTGAATTTGTCTCCATCTGGATCCTGTCTTATCTCTGGC 1134
 |||||
 201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
 |||||
 1135 CATTCCTGAAGCGATTTGGCTCTCTGATCTGCTGACCTTTGAATATAGGGTG 1184
 |||||
 217 alleProGluAlaIleGlyPheValMetValProPheGluTyrArgGlyG 234
 |||||
 1185 AACAGCATAAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
 |||||
 234 LuGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
 |||||
 1235 TACCAAGATGTAAGACCTGGTGGCTCTCGGCTCTTATTTCTGTATGCC 1284
 |||||
 251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPr 267
 |||||
 1285 CTTGGTGTCACTGCGATCTCTTACACCTCTCATGACTGTGTGAGATTTGA 1334
 |||||
 267 OleuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
 |||||
 1335 ACAGAAGGAATGCAGCTTCAGAAATTCCTCAGTGAACATCTTAAGCAG 1384
 |||||
 284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
 |||||
 1385 CTTGAGAAGTGGCAAAACAGTTTCTGCTGGTGTGTAATTTTGTCTCT 1434
 |||||
 301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
 |||||
 1435 TTGCTGGTTCCTCTTCACTTAAGCGTATATTGAAGAAACATCTGTCTATA 1484
 |||||
 317 ucYsrPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
 |||||
 1485 ACGAAATGGACAAGAACCGATGTGAATTAATTACTTACTTCTTACTGCTCATG 1534
 |||||
 334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
 |||||
 1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATTAACCCCAT 1584
 |||||
 351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProI 367
 |||||
 1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTCTTCCAGTCAATGCC 1634
 |||||
 367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
 |||||
 1635 TGTGCTGCTGTGTTTACCAGTCCAAAAGTCTGATGACCTCGGTGCCCATG 1684
 |||||
 384 euCysCysCysTyrGlnSerLysSerLeuMetThrSerValProMet 400
 |||||
 1685 AACGGAACAGCATCCAGTGGGAAGAACCCAGATCAAAACACCAACAC 1734
 |||||
 401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnAsnHisAsnTh 417
 |||||
 1735 AGCCGGAGCAGCCATAAGGACGATGAAC 1765
 |||||
 417 rAspArgSerSerHisLysAspSerMetAsn 427
 |||||

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.AAR53078

seq_documentation_block:

ID AAR53078 standard; Protein; 427 AA.
 AC AAR53078;
 DT 13-DEC-1994 (first entry)
 XX Human endothelin receptor subtype A (ETA).
 DE Endothelin; receptor; ETA; identification; detection; screening;
 KW inhibition.
 XX Homo sapiens.
 OS WO9409116-A.
 PN 28-APR-1994.
 XX 14-OCT-1993; 93WO-US09882.
 PF 15-OCT-1992; 92US-0961555.
 PR 15-OCT-1992; 92US-0961568.
 XX (MERI) MERCK & CO INC.
 PA Alves K, Chan CP, Hollis GF, Jones K, Tung J, Williams DL;
 PI WPI; 1994-151302/18.
 XX N-PSDB; AAQ63209.
 DR Transformed chinese hamster ovary cells - expressing human
 XX endothelin receptor subtypes A and B
 PT Example 1; Figure 2a-2c; 17pp; English.
 PS The human endothelin receptor subtype A (ETA) coding sequence may be
 CC used to transform chinese hamster ovary (CHO) cells. These cells
 CC then express ETA on their surface and can be used to identify
 CC substances which specifically inhibit the binding of endothelin to
 CC its receptor. Such an inhibitor could be used to treat patients with
 CC endothelin-related pathogenesis.
 XX
 SQ Sequence 427 AA;

alignment_scores:
 Quality: 2282.00 Length: 427
 Ratio: 5.344 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:
 US-09-931-157-1 x AAR53078 ...

Align seg 1/1 to: AAR53078 from: 1 to: 427

485 ATGGAACCCCTTGCTCAGGCATCCTTTGGCTGGCAGCTGGTGGATG 534
 |||||
 1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyC 17
 535 TGTATCACTGATATCCTCGAGAGATACAGCACAAATCTAAGCAATCATG 584
 |||||
 17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisV 34
 585 TGGATCATTTTACCACCTTTTGGGCACAGAGCTCAGCTTCCTGGTTACC 634
 |||||
 34 aAlaSerPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
 635 ACTCATCAACCCCACTAATTGGTCTTACCAGCAATGGCTCAATGCACAA 684
 |||||
 51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMethisAs 67
 685 CTATTGCCACAGCAGCTAAATTTACTTCAGCTTTCAAAATACATTAAACA 734
 |||||
 67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnT 84

735 CTGTGATATCTTGTAATACTATTTTCATCGTGGGAATGGTGGGAATGCAACT 784
 |||||
 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
 785 CTGCTCAGGATCATTACAGAACAAATGATGAGGAATGGCCCAACGC 834
 |||||
 101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
 835 GCTGATAGCCAGCTTGGCCCTTGGAGACCTTATCTATGTGGTCATTGATC 884
 |||||
 117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
 885 TCCCTATCAATGATATTAAAGTCGTGGCTGGCGCTGGCCTTTTGATCAC 934
 |||||
 134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHis 150
 935 AATGACTTTGGCGTATTTCTTTGCAAGCTGTTCCCTCTTTTCAGAGATC 984
 |||||
 151 AsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSe 167
 985 CTCGGTGGGATCACCGTCTCTCAACCTCTGCGCTCTTAGTGTGGACAGT 1034
 |||||
 167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
 1035 ACAGAGCAGTTCCCTCGGAGTCGTGTCAGGGAATCGGATTCCTTTG 1084
 |||||
 184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
 1085 TTAACCTGCCATTGAAATTCCTCCATCTCGATCGCTCCTCTCTTATCTGCG 1134
 |||||
 201 valThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
 1135 CATTCCTGAAGGATTTGGCTCTCATGTACCTTTGAATATAGGGGTG 1184
 |||||
 217 alleProGluAlaIleGlyPheValMetValProPheGluTyrArgGly 234
 1185 ACAGCATAAACCTGTATGCTCAATGCCACATCAAAATCATGGAGTTC 1234
 234 luGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
 1235 TACCAAGATGTAAGGACTGGTGGCTCTTCGGGTCTCTATTCTGTATGCC 1284
 |||||
 251 TyrGlnAspValLysAspTrpPheLeuPheGlyPheTyrPheCysMetPr 267
 1285 CTTGTGTGCACTGCGATCTTCTACACCTCATGACTTGTGAGATGTGA 1334
 |||||
 267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
 1335 ACAGAGGAATGGCAGCTTGAGAAATTCGCTCAGTGAACATCTTAAGCAG 1384
 |||||
 284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
 1385 CGTCGAGAAGTGCAGAAACAGTTTCTGCTGTTGGTTGAATTTTCTCTCT 1434
 |||||
 301 ArgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLe 317
 1435 TTGCTGTTCCCTCTTCACTTAAGCGGTATATTGAAGAAAACCTGTGTATA 1484
 |||||
 317 uCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
 1485 ACGAAATGGACAAGAACCGATGTGAATTAATTAGTTTCTTACTGCTCATG 1534
 |||||
 334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuMet 350
 1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAACCCCAT 1584
 |||||
 351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIl 367
 1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTGTTCCAGTCAATGCC 1634
 |||||
 367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384

1635 TCTGCTGCTGCTGTACCAAGTCTGATGACCTCGGTCCCATG 1684
 |||||
 384 eucyscyscysyrGlnSerLysSerLeuMetThrSerValProMet 400
 |||||
 1685 AACGGAACAAGCATCCAGTGGGAAGAACACGATCAAAACACCAACAC 1734
 |||||
 401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnAsnHisAsnTh 417
 |||||
 1735 AGACGGAGAGCCATAGGACAGCATGAAC 1765
 |||||
 417 rAspArgSerSerHisLysAspSerMetAsn 427

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.ABB56351

seq_documentation_block:

ID ABB56351 standard; Protein: 427 AA.

XX

AC ABB56351;

XX

DT 18-FEB-2002 (first entry)

XX

DE Non-endogenous human GPCR protein, SEQ ID NO: 495.

XX

KW Human; G protein-coupled receptor; GPCR: non-endogenous; mutant; constitutively activated GPCR; agonist; disease.

XX

OS Homo sapiens.

OS Synthetic.

XX

PN WO200177172-A2.

XX

PD 18-OCT-2001.

XX

PF 05-APR-2001; 2001WO-US11098.

XX

PR 07-APR-2000; 2000US-195747P.

XX

PA (AREN-) ARENA PHARM INC.

XX

PI Lehmann-Bruinsma K, Liaw CW, Lin I;

XX

DR WPI; 2001-648759/74.

XX

DR N-PSDB; AB197987.

XX

PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in disease treatment, comprises contacting candidate compounds with versions of GPCRs.

XX

PS Claim 1; Page 295-297; 394pp; English.

XX

CC The invention relates to G protein-coupled receptors (GPCRs) for which the endogenous ligand has been identified. Non-endogenous CC constitutively activated versions of known GPCRs are used in the CC invention for the direct identification of candidate compounds as CC receptor agonists, inverse agonists or partial agonists. Such CC agonists are useful as therapeutic agents for diseases or disorders CC associated with GPCRs. The present sequence is a non-endogenous CC version of a known human GPCR.

XX

SQ Sequence 427 AA;

alignment_scores:

Quality: 2277.00

Ratio: 5.345

Length: 427

Gaps: 0

Percent Similarity: 99.766

Percent Identity: 99.766

alignment_block:

US-09-931-157-1 x ABB56351

Align seq 1/1 to: ABB56351 from: 1 to: 427

485 ATGGAACCCCTTGGCTCAGGGCATCCTTTTGGCTGCGACTGGTGGATG 534

|||||
 1 MetGluThrLeuCysLeuArgAlaSerPheTrpLeuAlaLeuValGlyCy 17
 |||||
 535 TCTAATCAGTGATAATCCTGAGAGATACAGCACAAATCTAAGCAATCATG 584
 |||||
 17 sValIleSerAspAsnProGluArgTyrSerThrAsnLeuSerAsnHisv 34
 |||||
 585 TGGATGATTTCACCACTTTTTCGTGGCACAGAGCTCAGCTTCCTGGTTACC 634
 |||||
 34 aLaspAspPheThrThrPheArgGlyThrGluLeuSerPheLeuValThr 50
 |||||
 635 ACTCATCAACCCCACTAATTTGGTCTTACCCAGCAATGGCTCAATGCACAA 684
 |||||
 51 ThrHisGlnProThrAsnLeuValLeuProSerAsnGlySerMetHisas 67
 |||||
 685 CTATTGCCCCACAGCAGACTAAAATTTACTTTCAGCTTTCAAATACATAACA 734
 |||||
 67 nTyrCysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsn 84
 |||||
 735 CTGTGATATCTTGTACTATTTCATCGTGGGAATGGTGGGAATGCAACT 784
 |||||
 84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
 |||||
 785 CTGCTCAGGATCATTTACAGAACAAATGTATGAGGAATGGCCCCAACGC 834
 |||||
 101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
 |||||
 835 GCTGATAGCCAGCTCTGCCCTTGGAGACCTTATCTATGTGGTCATTGATC 884
 |||||
 117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValIleAspL 134
 |||||
 885 TCCCTATCAATGATTATTAACTGCTGGCTGGCGCTGGCTTTTGTATGCAC 934
 |||||
 134 euProIleAsnValPheLysLeuLeuAlaGlyArgTrpPropheAspHis 150
 |||||
 935 AATGACTTTGGCGTATTCTTTTGCAGCTGTTCCTCCCTTTTGCAGAGTC 984
 |||||
 151 AsnAspPheGlyValPheLeuLeuLysLysLeuPhePropheLeuGlnLysSe 167
 |||||
 985 CTCGTGGGGATCACCGCTCCTCAACCTCTGCGCTCTTAGTGTGACAGGT 1034
 |||||
 167 rSerValGlyIleThrValLeuAsnLeuLysLysLysLysLysLysLysLys 184
 |||||
 1035 ACAGAGCAGTTGCCCTCCCTGAGTCTGTCAGGGAATGGGATTCCTTTG 1084
 |||||
 184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
 |||||
 1085 GTAAGTGCATTAATGTCCTCCATCTGATCTGATCTGCTCTTATCTGCTGGC 1134
 |||||
 201 ValThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217
 |||||
 1135 CATTCCTCAAGCGATTGCTGCTGATGTCACCTTGTGAATATAGGGTG 1184
 |||||
 217 alleProGluAlaIleGlyPheValMetValProPheGluTyrArgGlyG 234
 |||||
 1185 AACAGCATAAACCTGTATGCTCAATGCCACATCAAAATTCATGGAGTTC 1234
 |||||
 234 luGlnHisLysThrCysMetLeuAsnAlaThrSerLysPheMetGluPhe 250
 |||||
 1235 TACCAAGATGTAAAGGACTGGTGGCTCTTCGGGTCTTATTTCTGTATGCC 1284
 |||||
 251 TyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPr 267
 |||||
 1285 CTTGGTGTGCACTGGGATCTTCTACACCTCATGACTTGTGAGATGTGA 1334
 |||||
 267 oLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 284
 |||||
 1335 ACAGAAGGAATGGCAGCTTGAGAATTCCTCCCTCAGTGAACATCTTAAGCAG 1384
 |||||
 284 snArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGln 300
 |||||
 1385 CGTCGAGAAGTGGCAAAACAGTTTCTGCTTGGTTGTAAATTTTCTCTCT 1434
 |||||

301 ArgArgGluValLysLysThrValPheCysLeuValValIlePheAlaLe 317
1435 TTGCTGGTTCCTCTTCACTTAAGCCGTATATTGAAGAAACTGTGTATA 1484
|||||
317 ucysrrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrA 334
1485 ACGAATGGACAAGAACCGATGTGAATTTACTTTAGTTTCTTACTGCTCATG 1534
|||||
334 snGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMet 350
1535 GATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATATAACCCCAT 1584
XX
351 AspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIle 367
1585 AGCTCTGTATTTTGTGAGCAAGAAATTTAAAAATTTTCCAGTCATGCC 1634
|||||
367 eAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCysL 384
1635 TCTGCTGCTGCTTACCAGTCCAAAGTCTGATGACCTCGGTGCCCATG 1684
384 euCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValPromet 400
1685 AACGGAAACAGCATCCAGTGGAGAACACGATCAAAACAAACACCAACAC 1734
|||||
401 AsnGlyThrSerIleGlnTrpLysAsnHisAspGlnAsnAsnHisAsnTh 417
1735 AGACCCGAGCAGCATTAAGCAGACATGAAC 1765
417 rAspArgSerSerHisLysAspSerMetAsn 427
seq_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:AA48719
seq_documentation_block:
ID_AA48719 standard; Protein; 333 AA.
XX
AC_AA48719;
XX
DT_06-JUN-1996 (first entry)
XX
XX G-protein coupled cow endothelin 1 receptor protein.
XX
XX G-protein coupled receptor; ligand binding assay; transmembrane domain;
KW psychotrophic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;
KW muscarinic acetylcholine; adrennergic; endothelin; bombesin; endocrine;
KW rhodopsin; opsin; odorant; cytomegalovirus.
XX
OS Bos taurus.
XX
XX WO9405695-A1.
PN
XX
PD 17-MAR-1994.
XX
PF 09-SEP-1993; 93WO-US08528.
XX
PR 10-SEP-1992; 92US-0943236.
XX
PA (UYNV) UNIV NEW YORK STATE.
XX
PI Murphy RB, Schuster DI;
XX
XX WPI; 1994-101120/12.
XX
XX Polypeptides of G-coupled receptor proteins (GPRs) - useful for
PT binding GPR ligands or modulating GPR binding
XX
PS Disclosure; Page 94-95; 160pp; English.
XX
XX Proteins AAR48685-R48758 represent a range of G-protein coupled receptor
CC proteins selected from cAMP, adenosine, muscarinic acetylcholine,
CC adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin,
CC odorant, cytomegaloviral and other G-protein coupled receptors. The
CC receptor proteins were used to design polypeptides, pref. based on the
CC transmembrane domains, for use in G-protein coupled receptor ligand

CC binding assays. The polypeptide fragments retain biological activity
CC such as binding a GPR ligand or modulating GPR ligand binding to a GPR
CC (see AAR48759-R48758, AAR50569-R50807 and AAR89189-R89195 for examples
CC of polypeptide fragments). The polypeptide fragments can be used in
CC compositions for treating subjects suffering from a pathology related to
CC a GPR abnormality e.g. a psychotic disorder such as schizophrenia.
XX
SQ Sequence 333 AA;

alignment_scores:
Quality: 1571.00 Length: 347
Ratio: 4.864 Gaps: 5
Percent Similarity: 93.084 Percent Identity: 85.303

alignment_block:
US-09-931-157-1 x AAR48719 ..

Align seg 1/1 to: AAR48719 from: 1 to: 333

725 TACATTACACTGTGATATCTTGTACTATTTTCATCTGGGATGCTGGG 774
|||||
1 TyrIleAsnThrIleValSerCysLeuValPheValLeuGlyIleIleG 17

775 GAATGCAACTCTCTCAGCATCATTTTACCAGAACAAATGTATGAGGAATG 824
|||||
17 yAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArgAsnG 34

825 GCCCAACCGCTGATAGCCAGTCTTGGCCCTTGGAGACCTTATCTATGTG 874
|||||
34 lProAsnIleLeuIleAlaSerIleAlaLeuGlyAspLeuLeuHisIle 50

875 GTCATGTATCCCTATCAATGTATTTAAGCTGCTGGCTGGCGCTGGCC 924
|||||
51 IleIleAspIleProIleMetAlaTyrLysLeuIleAlaGlyAspTrpPr 67

925 TTTTGATCACAATGACTTTGGCGTATTTCTTTCGAAGCTGTTCCTCTTT 974
|||||
67 o.....PheAlaCysLysLeuPheProPheL 76

975 TGCAGAAAGTCTCGTGGGATCACCGTCTCTCAACCTCTGCGCTCTTAGT 1024
|||||
76 euGlnLysSerSerValGlyIleThrValLeuAsnLeuCysAlaLeuSer 92

1025 GTTGACAGGTACAGACAGTTCCTCTCTGGAGTCTGTTTCAGGGAATTGG 1074
|||||
93 ValAspArgTyrArgAlaValAlaSerTrpSerArgValGlnGlyIleG 109

1075 GATTCTCTTGGTAACCTGCCATTCGAATTTGTCCTCATCTGGATCTGTCCT 1124
|||||
109 yIleProLeuValThrAlaIleGluIleValSerIleIleTrpIleLeuSerP 126

1125 TTATCTGCGCATTCCTGAAGCGATTGGCTTCCTCATGTGTACCTTTGAA 1174
|||||
126 heIleLeuAlaIleProGluAlaIleGlyPheTrpMetValProPheGlu 142

1175 TATAGGGTGAACAGCATAAACCTGTATCTCTCAATGCCACATCAAAATT 1324
|||||
143 TyrLysGlyAlaGlnHisArgThrCysMetLeuAsnAlaThrSerLysLe 159

1225 CATGGAGTTCTACCAAGATGTAAAGGACTGTGGCTCTCTCGGTTCTATT 1274
|||||
159 u.....PheTyrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrP 174

1275 TCTGTATGCCCTTGGTGTGCACCTGCGATCTTCTACACCTCATGACTGT 1324
|||||
174 heLeu.....LeuValCysThrAlaIlePheTyrThrLeuMetThrCys 188

1325 GAGATGTGTAACAGAAGAAATGGCAGCTTGAGAAATTCGCCCTCAGTGAACA 1374
|||||
189 GluMetLeuAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHi 205

1375 TCTTAAGCAGCGTCGAGAAAGTGGCAAAACAGTTTCTGCTGTTGTTAA 1424

250 oLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMetGlnP 267
1233 TCACCAAGATGTAAGACGTGGGTCTTCGGGTTCTATTTCGTATG 1282
267 heTyLysThrAlaLysAspTrpLeuPheSerPheTyPheCysLeu 283
1283 CCCTTGCTGTCACCTGCAGTCTTACACCTCATGACTGTGAGATGTT 1332
284 ProLeuAlaIleThrAlaIlePheTyThrLeuMetThrCysGluMetLe 300
1333 GAACAGAAAGTGGCAGTGTGAGAAATGCCTCAGTGAACATCTTAAGC 1382
300 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 316
1383 AGCGTCAGAAAGTGGCAAAACAGTTTCTGCTTGGTTGTAATTTTGCT 1432
316 lnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 332
1433 CTTTGTGCTTCCCTCTTCTACCTTAAGCCGTATATTGAAGAAACTGTGTA 1482
333 LeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTy 349
1483 TAACGAAATGGACAAGACCGATGTGAATTAATTAATTTTCTTACTGCTCA 1532
349 rAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeuVal 366
1533 TGATTACATCGGTATTAATCTTGGCAACCATGAATTCATCTATATAACCCC 1582
366 euAspTrpIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 382
1583 ATAGCTCTCTATTGTCAGCAAGAAATTTAAATTTGTTCCAGCATG 1632
383 IleAlaLeuTyThrValSerLysArgPheLysAsnCysPheLysSerCy 399
1633 CCTCTGCTGCTGCTGTACAGTCCAAAGTCTGATGACCTCGTCCCCA 1682
399 sleuCysCysTrpCys...GlnThrPheGluGluLysGlnSerLeuGluG 415
1683 TGAACGGAACAGCATCCAGTGGAGAACACCATCAACAAACACCAAC 1732
415 LuLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyAsp 430
1733 ACAGACCGGAGCAGGCATTAAGACAGC 1759
431 AsnPheArgSerSerAsnLysTySer 439

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.AAR30886

seq_documentation_block:
ID AAR30886 standard; Protein; 442 AA.

XX AAR30886;
AC
XX
DT 11-MAY-1993 (first entry)
XX
DE ETb receptor.
XX
KW Human; ETa; ETb; endothelin; receptor; transmembrane domain; N tail;
KW extracellular; cytoplasmic; C tail; post translational; bovine;
KW modification; ET-1 receptor; antagonist; circulatory system.
XX
OS Homo sapiens.
XX
FH Key
FT Peptide
FT
FT Protein
FT
FT Domain
FT Domain
FT
FT
FT
FT
FT

Location/Qualifiers
1..26 /note= "Signal peptide"
27..442 /note= "Mature peptide"
102..126 /note= "Transmembrane domain I"
138..163 /note= "Transmembrane domain II"

Domain 176..197
/note= "Transmembrane domain III"
Domain 219..243
/note= "Transmembrane domain IV"
Domain 272..296
/note= "Transmembrane domain V"
Domain 325..350
/note= "Transmembrane domain VI"
Domain 363..389
/note= "Transmembrane domain VII"
Modified-site 59
/note= "N-glycosylated site"
Modified-site 305
/note= "May be phosphorylated with serine/threonine
kinases"
Modified-site 413
/note= "May be phosphorylated with serine/threonine
kinases"
Modified-site 419
/note= "May be phosphorylated with serine/threonine
kinases"
Modified-site 435
/note= "May be phosphorylated with serine/threonine
kinases"
Modified-site 436
/note= "May be phosphorylated with serine/threonine
kinases"
Modified-site 440
/note= "May be phosphorylated with serine/threonine
kinases"
EP522868-A.
PN
XX
XX
PD 13-JAN-1993.
XX
PF 10-JUL-1992; 92EP-0306347.
XX
PR 12-JUL-1991; 91JP-0172828.
XX
PA (SHIO) SHIONOGI SEIYAKU KK.
XX
PI Imura H, Nakanishi S, Nakao K;
XX
DR WPI; 1993-010677/02.
XX
DR N-PSDB:AAQ34584.
XX
PT Human ETa and ETb endothelin receptors - for measuring endothelin
PT and screening for endothelin antagonists
PS Claim 8; Fig 2; 39pp; English.
XX
CC The sequences given in AAR30885-86 represent the human ETa and ETb
CC endothelin receptors respectively. ETa is a 427 amino acid protein
CC with a molecular weight of 48,726. ETb comprises 442 amino acids and
CC has a molecular weight of 49,629. ETa has a higher affinity for
CC endothelin (ET)-1 and ET-2, whereas ETb has no selectivity for ET-1,
CC ET-2 or ET-3. The receptors each contain seven transmembrane domains
CC and have an extracellular N tail and a cytoplasmic C tail. There are
CC several potential sites for post translational modification, these
CC sites are identical to those of bovine ET-1 receptor. ETa cDNA is
CC 91.2% homologous to bovine ET-1 receptor cDNA and ETb cDNA is 61.1%
CC homologous to that of bovine ETa-receptor. The receptor proteins are
CC useful as reagents for measuring the amount of ET or screening for
CC antagonists of the ET receptor when studying the circulatory system.
XX
SQ Sequence 442 AA;

alignment_scores:
Quality: 1240.00 Length: 359
Ratio: 3.987 Gaps: 5
Percent Similarity: 86.630 Percent Identity: 63.510

Percent Similarity: 86.351 Percent Identity: 63.231

alignment_block:

US-09-931-157-1 x ABB56352 ..

Align seg 1/1 to: ABB56352 from: 1 to: 442

```
689 TGCCACAGCAGACTAAATTTACTTTCAGCTTCAATACATTAACACTGT 738
|||||
90 CysGlnGlyProIleGluIleLysGluThrPheLysThrIleAsnThrVa 106
739 GATATCTTGACTATTTCATCGTGGAAATGGTGGGAATGCAACTCTGC 788
|||||
106 lValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeuL 123
789 TCAGGATCAATTTACCAAGAACAAATATGATGAGGAATGCCCAACGCGCTG 838
|||||
123 euArgIleIleTyLysAsnLysCysMetArgAsnGlyProAsnIleLeu 139
839 ATAGCAGCTCTTCCCTTGGAGACCTTATCTATGTGCTATTGATCTCCC 888
|||||
140 IleAlaSerLeuAlaLeuGlyAspLeuLeuHisIleValIleAspIlePr 156
889 TATCAATGATTTAAGCTGCTGGCGCTGGCGCTGCTTTTGTATCACATG 938
|||||
156 OrlaenValTyLysLeuLeuAlaGluAspTrpPro..... 168
939 ACTTTGGCGTATTTCTTCAAGCTCTTCCCTTTTGGCAGAAGTCTCG 988
|||||
169 ..PheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAlaSer 184
989 GTGGGATCACCGTCTCAACCTCTCGCTCTTGTAGTTGACAGGTACAG 1038
|||||
185 ValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyRar 201
1039 ACAGTTGCTCTCTGAGTCGTGTCAGGAATGGGATTCCTTTGGTAA 1088
|||||
201 gAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpT 218
1089 CTGCAATGAAATTTCTCATCTGATCTGCTTCTTCTTCTTCTGCGCAT 1138
|||||
218 hrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVal 234
1139 CTGAGGAGTGGCTTCTGTCAGTACCTTTGATATAGGAGGTGAACA 1188
|||||
235 ProGluAlaIleGlyPheAspIleIleThrMetAspTyLysGlySerty 251
1189 GCATAAAACCTGTATGCTCAATGCCACATCAAAA.....TTCATGGAGT 1232
|||||
251 rLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetGlnP 268
1233 TCTACCAAGATGTAAGGACTGGTGGCTCTCGGCTTCTATTCTGTATG 1282
|||||
268 heTyLysThrAlaLysAspTrpLeuPheSerPheTyPheCysLeu 284
1283 CCCTGTGTGCATCGATCTTCTACACCTCATGACTGTGAGATGT 1332
|||||
285 ProLeuAlaIleThrAlaPhePheTyThrLeuMetThrCysGluMetLe 301
1333 GAACAGGAAGTGGCAGCTTGAGATTCCTCAGTCAATCTTAAGC 1382
|||||
301 uArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLysG 317
1383 AGCTCGAGAAGTGGCAAAACAGTTTCTGCTGGTGTAAATTTTCTGCT 1432
|||||
317 lnArgArgGluValLysLysThrValPheCysLeuValLeuValPheAla 333
1433 CTTTGTGTTCCCTCTTCACTTAAGCGGTATATTGAAGAAACTGTGTA 1482
|||||
334 LeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTy 350
1483 TAAGAAATGGACAGAAGACCGATGTGAATTAATTAATTAATTAATTAAT 1532
|||||
```

```
350 rAsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuValL 367
1533 TGGATTACATCGGTATTAACTTGGCAACCATGAATTCATGTATAAACCCC 1582
|||||
367 euAspTyRileGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 383
1583 ATAGCTCTGTATTTTGTGAGCAAGAAATTTAAATTTGTTTCCAGTCATG 1632
|||||
384 IleAlaLeuTyRLeuValSerLysArgPheLysAsnCysPheLysSerCy 400
1633 CCTGTCTGCTGCTTACCAAGTCCAAAGTCTGATGACCTCGGTCCCA 1682
|||||
400 sLeuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGluG 416
1683 TGAACGGAACAGCATCCAGTGGAGAACACACGATCAAAACACACACAC 1732
|||||
416 luLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyRAsp 431
1733 ACAGACCGGAGCAGCCATTAAGCAGC 1759
|||||
432 AsnPheArgSerSerAsnLysTyRSer 440
```

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:AAW59436

seq_documentation_block:

ID AAW59436 standard; protein; 452 AA.

AC AAW59436;

DT 27-AUG-1998 (first entry)

DE Human endothelin-beta 1 receptor protein.

KW Endothelin beta-1 receptor; ETB-1; screening; signal transduction;
KW agonist; antagonist; vasoconstrictor; vasopressor; cardiogenic shock;
KW pulmonary hypertension; acute myocardial infarct; uraemia; vasculitis;
KW Crohn's disease; ulcerative colitis; sepsis; congestive heart failure;
KW coronary spasm; cyclosporin nephrotoxicity; toxemia.

OS Homo sapiens.

FX Key Location/Qualifiers
FT Misc-difference 199..208 /label- decapeptide insert
XX US5773223-A.

XX 30-JUN-1998.

XX 02-SEP-1993; 93US-0117361.

XX 02-SEP-1993; 93US-0117361.

XX (CHIR) CHIRON CORP.

XX Olson PT, Shyamala V;

XX WPI; 1998-386992/33.

XX Screening for modulators of the endothelin B1 receptor - by
PT measuring effect on signal transduction in cells engineered to
PT express the receptor, potentially useful as agonists and antagonists
PT of endothelin

XX Claim 1a; Column 31-34; 23pp; English.

CC This sequence represents a novel human endothelin-beta1 receptor (ETB-1).
CC This sequence is used in a method involving the screening of compounds
CC for their ability to bind to endothelin B1 (ETB1) receptor polypeptide
CC and to modulate its signal transduction activity by applying test
CC compound to host cells transfected with DNA encoding ETB-1 and optionally
CC lysing the cells and then measuring signal transduction activity. The
CC method is used to identify agonists and antagonists of endothelin (ET), a

XX Bovine endothelin receptor sub-type B - used for elucidation of
 PT mechanism of ischaemic tissue interference and its treatment,
 PT diagnosis and prevention
 XX
 PS Claim 1; Fig 1; 9pp; Japanese.
 XX
 CC This sequence represents the bovine endothelin (ET) receptor subtype
 CC 2. This receptor protein may be used for the elucidation of the
 CC mechanism of ischaemic tissue interference and its treatment,
 CC diagnosis and prevention. The cDNA encoding this sequence was
 CC isolated from bovine lung tissue.
 XX
 SQ Sequence 441 AA;

alignment_scores:
 Quality: 1209.50 Length: 376
 Ratio: 3.840 Gaps: 6
 Percent Similarity: 83.777 Percent Identity: 59.574
 alignment_block:
 US-09-931-157-1 x AAR43992
 Align seg 1/1 to: AAR43992 from: 1 to: 441
 659 CTACCCAGCAATGGCTCAATGCACAACTAT..... 688
 72 IleProArgGlyGlyArgMetAlaGlyLeProProArgThrProPro 88
 689 .TCCCCACAGCAGCACTAAATTTACTTCAGCTTCAATATACATTAACACTG 737
 88 oCysAspGlyProIleGluLeuIleValPheLysThrValLeuAsnThrV 105
 738 TGATATCTTGTTACTATTTCATCGTGGGAATGTGGGAATGCACTCTG 787
 105 alValSerCysLeuValPheValLeuGlyValLeuLeuHisIleIleAspIleP 121
 788 CTCAGGATCATTTACAGAACAAATATGAGGAATGGCCCAACGGCT 837
 122 LeuArgIleIleThrLysAsnLysCysMetArgAsnGlyProAsnIleLe 138
 838 GATAGCCACTCTGCCCTGGAGACCTTATCTATGATGGTCATGATCTCC 887
 138 uIleAlaSerLeuAlaLeuGlyValLeuLeuHisIleIleAspIleP 155
 888 CTATCAATGATTTAAGCTGTGGCTGGCGCTGGCTTTTGTATCACAAT 937
 155 toIleAsnThrTyrlLysLeuLeuAlaLysAspTrpPro..... 167
 938 GACTTTGGCGTATTTCTTTGCAAGCTGTTCCTCTTTTTCGACGAAGTCCTC 987
 168 ...PheGlyValGluMetCysLysLeuValProPheIleGlnLysAlaSe 183
 988 GTGGGGATCACCCTCTCACTCTGGCTCTTATAGTTGTACAGGTACA 1037
 183 rValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyra 200
 1038 GAGCAGTTGCTCTCTGGAGTGTCTTCAGGAATGGGATTCCTTTGGTA 1087
 200 rGAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrp 216
 1088 ACTGCCATTGAAATGTCCTCCATCTGGATCTGCTTTATCTTCCTGGCCAT 1137
 217 ThrAlaValGluIleLeuLeuIleThrValSerValValLeuAlaVal 233
 1138 TCCTGAAGCATTGGCTTCGTGATACCTCTTGAATATAGGGGTGAAC 1187
 233 lProGluAlaValGlyPheAspIleIleThrSerAspHisIleGlyAsnL 250
 1188 AGCATAAACCTGTATCTCAATGCCACATCAAAA.....TTCATGGAG 1231
 250 ysLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMetGln 266

1232 TTCTACCAAGATTAAGGACTGTGCTCTCTCGGGTCTATTCTGTAT 1281
 267 PheTyrlLysThrAlaLysAspTrpTrpLeuPheSerPheTyrlPheCysLe 283
 1282 GCCCTTGGTGTGACGCGCATCTTACACCCCTCATGACTTGTGAGATGT 1331
 283 uProLeuAlaIleThrAlaLeuPheTyrlThrLeuMetThrCysGluMetL 300
 1332 TGAACAGAAGGAATGCGACGCTTGAGAAATTCCTCAGTGAACATCTTAAG 1381
 300 euArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeuLys 315
 1382 CAGCGTCGAGAAGTGGCAAAACAGTTTCTGCTGTTGTTTATTTTTCG 1431
 316 GlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAl 332
 1432 TCTTTGCTGGTTCCTCTTCACTTAAGCGGTATATTGAAGAAACCTGTGT 1481
 332 aLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuT 349
 1482 ATAAGGAAATGCACAGAACCGATGTGAATTTACTTAGTTTCTTACTGCTC 1531
 349 yrAspGlnHisAspProArgArgCysGluPheLeuSerPheLeuLeuVal 365
 1532 ATGATTACATCGGTATTAACTTTGGCAACCATGAATTCATGTATAACCC 1581
 366 LeuAspTyrlleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPr 382
 1582 CATAGCTCTGTATTTTGTGACCAAGAAATTTAAAATTTGTTTCCAGTCAT 1631
 382 oIleAlaLeuTyrlLeuValSerLysArgPheLysAsnCysPheLysSerC 399
 1632 GCCCTCTGCTGCTGTACCAAGTCCAAACCTCATGACCTCGGTCCCC 1681
 399 ysLeuCysCysTrpCys...GlnSerPheGluGluLysGlnSerLeuGlu 414
 1682 ATGAACGAAGAACAGCATCCAGTGAAGAACCAACGATCAAAACACCAAA 1731
 415 GluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTyraS 430
 1732 CACAGACCGGAGCAGCCATAGGACAGC 1759
 430 pAsnPheArgSerSerAsnLysTyrlSer 439

seq_name: /SIDSL/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48720

seq_documentation_block:

ID AAR48720 standard; Protein; 350 AA.

AC AAR48720;

XX

XX 06-JUN-1996 (first entry)

DT

XX G-protein coupled rat non-isopeptide selective endothelin receptor.

DE

XX G-protein coupled receptor; ligand binding assay; transmembrane domain;

XX psychotic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;

XX muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;

XX rhodopsin; opsin; odorant; cytomegalovirus.

XX Rattus rattus.

OS

XX WO9405695-A1.

PN

XX 17-MAR-1994.

PD

XX 09-SEP-1993; 93WO-US08528.

PF

XX 10-SEP-1992; 92US-0943236.

PR

XX (UYN) UNIV NEW YORK STATE.

PA

XX

PI Murphy RB, Schuster DI;
 DR WPI; 1994-101120/12.
 XX Polypeptides of G-coupled receptor proteins (GPRs) - useful for
 PT binding GPR ligands or modulating GPR binding
 XX
 XX Disclosure; Page 95-96; 160pp; English.
 XX
 CC Proteins AAR4895-R48758 represent a range of G-protein coupled receptor
 CC proteins selected from cAMP, adenosine, muscarinic acetylcholine,
 CC adrenergic, thrombin, endothelin, bombesin, endocrine, opsin,
 CC odorant, cytomagaloviral and other G-protein coupled receptors. The
 CC receptor proteins were used to design polypeptides, pref. based on the
 CC transmembrane domains, for use in G-protein coupled receptor ligand
 CC binding assays. The polypeptide fragments retain biological activity
 CC such as binding a GPR ligand or modulating GPR ligand binding to a GPR
 CC (see AAR48759-R48758, AAR50569-R50807 and AAR89189-R89195 for examples
 CC of polypeptide fragments). The polypeptide fragments can be used in
 CC compositions for treating subjects suffering from a pathology related to
 CC a GPR abnormality e.g. a psychotic disorder such as schizophrenia.
 XX
 SQ Sequence 350 AA;

alignment_scores:
 Quality: 899.00 Length: 311
 Ratio: 3.596 Gaps: 7
 Percent Similarity: 80.386 Percent Identity: 57.235

alignment_block:

US-09-931-157-1 x AAR48720 ..

Align seg 1/1 to: AAR48720 from: 1 to: 350

865 TATCTATGTGGTCATTGATCTCCCTATCAATGATTTAAAGCTGCTG.... 910

 43 PheLeuGlyGlyLeuValLeuThrLeuAspPheLeuGlyLeuValMet 59
 911GCTGGGCGCTGGCTTTTGATCAC 934

 59 hrGlyThrIleValSerGlnHisAlaAlaLeuPheGluTrpHisAla 75
 935 AATGACTTTGGCGTATTTCTTTCGAAGCTGTTCCTCCCTTTTTCGAGAAGTC 984
 III III
 76 ValAspProGlyCysArgLeuValProPheIleGlnLysAl 92
 985 CTCGGTGGGATCACGCTCTCAACCTCTGCGCTCTAGTGTGACAGGT 1034

 92 aserValGlyIleThrValLeuSerLeuGlyCysAlaLeuSerIleAspArg 109
 1035 ACAGAGCAGTTCCTCTCGAGTCTGTTTCAGGGAAATGGGATTCCTTTG 1084

 109 yrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLys 125
 1085 GTAACGTGCCATTGAATGCTCCATCTCGATCCTCTCTTTATCTCTGCG 1134

 126 TrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuAl 142
 1135 CATTCCTGAAGGATGGCTTCGTCATCGTACCTTTTGAATATAGGGTG 1184

 142 ValProGluAlaIleGlyPheAspThrThrSer...AspTrpLysGlyL 158

 1185 AACAGCATAAACCTCTGATGCTCAATGCCACATCAAAA.....TTCATG 1228

 158 ysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMet 174
 1229 GAGTCTACACAGATGTAAGGACTGGTGGCTCTTCGGGTTCTATTCTG 1278

 175 PheTrpLysThrAlaAlaLysAspTrpTrpLeuPheAlaPheTrpPheCy 191

 1279 TATGCCCTGGTGTGACGCTGCTCTTACACCTCATGCTGTGAGA 1328

 1328

191 sleuProLeuAlaIleThrAlaIlePheTrpThrLeuMetThrCysGluM 208
 1329 TCTTGAACAGAGGAATGCAGCTTGAGAAATTCCTCAGTGAACATCTT 1378

 208 etLeuArgLysLysSerGly...MetGlnIleAlaLeuAsnAspHisLeu 223

 1379 AAGCAGCGTCGAGAAAGTGGCAAAACAGTTTCTGCTGCTGCTGCTGCT 1428

 224 LysGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPh 240

 1429 TGCTCTTTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1478

 240 ealaleuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrL 257

 1479 TGTATAACGAATGGACAAACCGATGTAATTAATTAATTAATTAATTAAT 1528

 257 eutyAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeu 273

 1529 CTCATGGATTACATCGGTATTAACTTGGCAACCATCAATTCATGTATAA 1578

 274 ValLeuAspTrpIleGlyIleAsnMetAlaSerIleAsnSerCysIleAs 290

 1579 CCCCATAGCTCTGATTTTGTGAGCAAGAAATTTAAATAATTTGTTCCAGT 1628

 290 nProIleAlaLeuTrpLeuValSerLysArgPheLysAsnCysPheLysS 307

 1629 CATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1678

 307 erCysLeuGlyCysTrpCys...GlnThrPheGluGlyLysGlnSerLeu 322

 1679 CCCATGAACGGAAACAGCATCCAGTGGAGAGAACCCAGCATCAAAACACCA 1728

 323 GluGluLysGlnSerCysLeuLysPheLysAlaAsnAsp...HisGlyTy 338

 1729 CAACAGACAGCGAGCAGCATCAAGGACAGC 1759

 338 rasPasnPheArgSerSerAsnLysTrpSer 348

 seq_name: /SIBS1/gcgdata/geneseq/geneseq-embl/AAL996.DAT:AAW02692
 seq_documentation_block:
 ID_AAW02692 standard; peptide; 350 AA.
 XX
 AC AAW02692;
 XX
 DT 12-NOV-1996 (first entry)
 XX
 DE G-protein coupled rat non-isopeptide selective endothelin receptor.
 XX
 KW G-protein coupled receptor; ligand binding assay; transmembrane domain;
 KW schizophrenia; dopamine; cAMP; adenosine; thrombin; adrenergic; opsin;
 KW muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin;
 KW odorant; cytomagalovirus; serotonergic.
 XX
 OS Rattus rattus.
 XX
 PN US5508384-A.
 XX
 PD 16-APR-1996.
 XX
 PF 10-SEP-1992; 92US-0943236.
 XX
 PR 09-SEP-1993; 93US-0118270.
 PR 10-SEP-1992; 92US-0943236.
 XX
 PA (UJNY) UNIV NEW YORK STATE.
 XX
 PI Murphy RB, Schuster DI;
 XX
 DR WPI; 1996-208785/21.
 XX
 XX New dopamine receptor peptide - useful as antipsychotic agent, e.g.
 PT for treating schizophrenia

PN WO9106647-A.
 XX 16-MAY-1991.
 PD 23-OCT-1990; 90WO-US06125.
 PF 23-OCT-1990; 90WO-US06125.
 PR 24-OCT-1989; 89US-0426150.
 PR 05-JUN-1990; 90US-0533659.
 XX (TRIT-) TRITON BIOSCIENCES.
 PA
 XX
 XX Feldman RI, Harkins RN, Batey J, Slattery TK;
 PI WPI; 1991-164196/22.
 XX N-PSDB; AAQ12032.
 DR
 XX DNA and polypeptide sequence of GRP receptor - used in treatment
 PT and diagnosis of cancerous tissues.
 PS Claim 5; Fig 8; 84pp; English.
 XX
 CC The purified receptor can be used in drug screening for cpds.
 CC having suitable binding affinity for GRPR. The receptor,
 CC fragments and antibodies may also be used in diagnostic assays
 CC to determine altered levels of GRPR. The antibodies are pref.
 CC raised to fragments comprising residues 1-39, 64-77, 98-115,
 CC 138-157, 176-209, 236-266, 288-300 and 330-385.
 XX
 XX Sequence 384 AA;
 SQ

 alignment_scores:
 Quality: 465.00 Length: 342
 Ratio: 2.076 Gaps: 7
 Percent Similarity: 65.497 Percent Identity: 32.456

 alignment_block:
 US-09-931-157-1 x AAR12399 ..
 Align seg 1/1 to: AAR12399 from: 1 to: 384

 719 TTCAATACATACACTGTGATATCTGTACTATTTTCATCGTGGGAAT 768
 ||| ||||| : : : : : ||||| : : : : :
 40 PheIleTyValIleProAlaValTyrglyLeuIleValIleGlyLe 56
 : : : : : ||||| : : : : : ||||| : : : : :
 769 GGTGGGAATGCACTCTGCTCAGCATCATTTACCAGAACAAATATGA 818
 : : : : : ||||| : : : : : ||||| : : : : :
 56 uileGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMeta 73
 ||||| ||||| : : : : : ||||| : : : : :
 819 GGAATGGCCCAACGCGCTGATAGCAGCTCTGCCCTTGGAGACCTTATC 868
 ||||| ||||| : : : : : ||||| : : : : :
 73 rGAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLeu 89
 : : : : : ||||| : : : : : ||||| : : : : :
 869 TATGTGTCATGATCTCCCTATCAATGATTTTAAAGCTGCTGGCTGGCG 918
 : : : : : ||||| : : : : : ||||| : : : : :
 90 LeuLeuValThrCysAlaProValAspAlaSerLysTyrrLeuAlaAspAr 106
 : : : : : ||||| : : : : : ||||| : : : : :
 919 CTGCCCTTTTGCATCACAATGACTTGGCGTATTCTTTGCAAGCTGTCC 968
 ||||| ||||| : : : : : ||||| : : : : :
 106 gtrPLeu.....PheGlyArgIleGlyCysLysLeuIleP 118
 : : : : : ||||| : : : : : ||||| : : : : :
 969 CCTTTTGCAGAGCTCTCGTGGGGATCAGCTCCTCAAGCTCTGCGCT 1018
 ||||| ||||| : : : : : ||||| : : : : :
 118 roPheIleGlnLeuThrSerValGlyValSerValPheThrLeuThrAla 134
 : : : : : ||||| : : : : : ||||| : : : : :
 1019 CTTAGTGTGACAGGTACAGACAGTGTCCCTCGAGTCTGTTTCAGGG 1068
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 135 LeuSerAlaAspArgTyrrLysAlaIleValArgPrometAspIleGlnAl 151
 : : : : : ||||| : : : : : ||||| : : : : :
 1069 AATTGGGATTCCTTGGTAAGTGCATTTGAATGCTCTCAATCTGGATCC 1118
 : : : : : ||||| : : : : : ||||| : : : : :
 151 aSerHisAlaLeuMetLysIleCysLeuLysAlaAlaLeuIleThrIleV 168

1119 TGTCTTTATCTCGCCATTCCTGAAGCGATTCGCTTCATCGTACCC 1168
 : : : : : ||||| : : : : : ||||| : : : : :
 168 alSerMetLeuAlaIleProGluAlaValPheSerAspLeuHisPro 184
 : : : : : ||||| : : : : : ||||| : : : : :
 1169 TTTGAATATAGGGTGAACACATATAAACCTGTATGCTCAATGCCACATC 1218
 ||||| : : : : : ||||| : : : : : ||||| : : : : :
 185 PheHisValLysAspThrAsnGlnThrPheIleSerCysAlaProTyrrPr 201
 : : : : : ||||| : : : : : ||||| : : : : :
 1219 AAAATTCATGGAGTCTTACCAAGATGTAAAGGACTGTGGCTCTTCGGGT 1268
 : : : : : ||||| : : : : : ||||| : : : : :
 201 oHisSerAsnGluLeuHisProLysIleHisSerMetAlaSerPheLeuV 218
 : : : : : ||||| : : : : : ||||| : : : : :
 1269 TCTATTTCTGTATGCTTGGTGTGCACATCGATCTTCTACACCTCATG 1318
 : : : : : ||||| : : : : : ||||| : : : : :
 218 alPheTyrrValIleProLeuAlaIleSerValTyrrTyrrPheIle 234
 : : : : : ||||| : : : : : ||||| : : : : :
 1319 ACTGTGTAGATGTGAACAGAGG.....AATGGCAGCTT 1353
 : : : : : ||||| : : : : : ||||| : : : : :
 235 AlaArgAsnLeuIleGlnSerAlaTyrrAsnLeuProValGluGlyAsnII 251
 : : : : : ||||| : : : : : ||||| : : : : :
 1354 GAGAAATGCCCTCAGTGAACATCTTAAGCAGCTCGAGAAAGTGCACAAA 1403
 : : : : : ||||| : : : : : ||||| : : : : :
 251 eHisVal.....LysLysGlnIleGluSerArgLysArgLeuAlaLysT 266
 : : : : : ||||| : : : : : ||||| : : : : :
 1404 CAGTTTCTGCTGGTGTGTAATTTTCTCTTCTTCTGCTGGTTCCTCTTCAC 1453
 : : : : : ||||| : : : : : ||||| : : : : :
 266 hrValLeuValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHis 282
 : : : : : ||||| : : : : : ||||| : : : : :
 1454 TTAAGCCGTATATTGAAGAAACTGTGTATAACGAAATGGAACAGACCG 1503
 : : : : : ||||| : : : : : ||||| : : : : :
 283 ValIleTyrrLeuTyrrArgSerTyrrHisTyrrSerGluValAspThrSer.. 298
 : : : : : ||||| : : : : : ||||| : : : : :
 1504 ATGTGAATTACTAGTTTCTTACTGCTCATCGATTACATCGGTATTAACT 1553
 : : : : : ||||| : : : : : ||||| : : : : :
 299MetLeuHisPhe.....ValThrSerIleCysAlaHisLeuL 311
 : : : : : ||||| : : : : : ||||| : : : : :
 1554 TGGCAACCATGATTCATGTATATAAACCCCATAGCTCTGTATTTTGTAGC 1603
 : : : : : ||||| : : : : : ||||| : : : : :
 311 euAlaPheThrAsnSerCysValAsnProPheAlaLeuTyrrLeuLeuSer 327
 : : : : : ||||| : : : : : ||||| : : : : :
 1604 AAGAAATTTAAAAATTTTCCAGTCATCGCTCTGCTGCTGTATTACCA 1653
 : : : : : ||||| : : : : : ||||| : : : : :
 328 LysSerPheArgLysGlnPheAsnThrGlnLeuLeuCysCys..... 341
 : : : : : ||||| : : : : : ||||| : : : : :
 1654 GTCCAAAGTCTGATGACCTCGGTCCCATGAACGGAACAAAGCATCCAGT 1703
 : : : : : ||||| : : : : : ||||| : : : : :
 342 .GlnProGlyLeuMetAsn.ArgSerHisSerThrGlyArgSerThrThr 357
 : : : : : ||||| : : : : : ||||| : : : : :
 1704 GGAAGAACCCAGCATCAAAACACC 1727
 : : : : : ||||| : : : : : ||||| : : : : :
 358 CysMetThrSerPheLysSerThr 365
 : : : : : ||||| : : : : : ||||| : : : : :
 seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:AA27507
 seq_documentation_block:
 ID: AAR27507 standard; Protein: 384 AA.
 XX
 AC AAR27507;
 XX
 DT 05-MAR-1993 (first entry)
 XX
 DE Mouse RIBP (GRP-R).
 XX
 KW Gastrin releasing peptide; GRP; bombesin; neuromedin B; NMB;
 KW ranatensin; RBP; bombesin-like peptide; RBP; R2BP; R2BP;
 KW receptor; agonist; antagonist; ligand; antibody; cancer.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Region 40..63
 FT /label= putative_transmembrane_sequence

1704 GGAAGAACCGATCAAAACAACC 1727

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 30, 2002, 20:31:22 : Search time 8437.74 Seconds
(without alignments)
6879.852 Million cell updates/sec

Title: us-09-931-157-2

Perfect score: 4301
Sequence: 1 gagacattccggtggggac.....ctggagaaaaa 4301

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST.*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_esti:*
10: gb_hst2:*
11: gb_hst:*
12: gb_gss:*
13: gb_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	937	21.8	960	9	AL571798
c 2	866	20.1	957	10	BI520706
c 3	842.6	19.6	874	9	AL543805
c 4	808.6	18.8	972	10	BI858627
c 5	798.6	18.6	972	9	AL546465
c 6	750.6	17.5	823	9	AL553041
c 7	746.6	17.4	796	9	AL553065
c 8	741.8	17.2	770	10	BM014035
c 9	729.2	17.0	758	10	BM014042
c 10	727.8	16.9	885	10	BG769122
c 11	725	16.9	842	9	AL571072
c 12	718	16.7	785	9	AU117045
c 13	712.8	16.6	743	9	AU138228
c 14	706.2	16.4	800	9	AU1760041
c 15	698.8	16.2	726	9	AL422064
c 16	698	16.2	885	9	AL545283
c 17	697.4	16.2	886	9	AI188458

18	696.8	16.2	771	9	AU116904
19	686.4	16.0	839	9	AU136164
c 20	679.4	15.8	771	9	AI567763
c 21	662.6	15.4	751	9	AA651686
c 22	659.6	15.3	726	9	AU134688
c 23	655.2	15.2	672	9	AA573116
c 24	647.4	15.1	752	9	AV710837
c 25	644.2	15.0	761	9	AL570142
c 26	643.6	15.0	969	10	BI757487
c 27	630.8	14.7	634	10	BE837728
c 28	625.8	14.6	629	9	AV646124
c 29	620	14.4	847	10	BI768770
c 30	617.8	14.4	756	10	BG620583
c 31	610.6	14.2	633	9	AU139933
c 32	610.6	14.2	631	9	AW665874
c 33	602.6	14.0	638	9	AI800586
c 34	599.6	13.9	744	10	BE622559
c 35	596.6	13.9	644	9	AW960695
c 36	591	13.7	670	9	AW192839
c 37	581.4	13.5	787	10	BG217179
c 38	578.6	13.5	605	10	BE827729
c 39	574	13.3	601	9	AI760190
c 40	562.2	13.1	637	9	AI808370
c 41	553	12.9	578	9	AI144515
c 42	541	12.6	563	9	AW469710
c 43	540.4	12.6	923	10	BG573693
c 44	539	12.5	555	10	BE837570
c 45	531.6	12.4	556	10	N32826

ALIGNMENTS

RESULT 1
AL571798/c
LOCUS AL571798 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI030YM19 3
DEFINITION prime, mRNA sequence.
ACCESSION AL571798
VERSION AL571798.1 GI:12929453
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 960)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..960
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI030YM19"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 293 a 221 c 198 g 246 t 2 others
ORIGIN

```
Query Match      21.8%; Score 937; DB 9; Length 960;
Best Local Similarity 99.0%; Pred. No. 5.8e-150;
Matches 952; Conservative 1; Mismatches 7; Indels 2; Gaps 1;

QY 559 tccctgcttgctgtctgtggtgatactcgggaactccacactcttgagaattattac 618
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
960 TCCCTGCTTGTGTCTGCTGGG--ATCATCGGRATCCACACTCTGAGAATTATCTAC 903

QY 619 aagaaacagtgcagcgaacggtcccaatatcttgatccagcttgctggagac 678
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
902 AAGAACAGTGCATGCGAAGCGTCCCAATATCTTGATCGCCAGCTGGCTCGGGAGAC 843

QY 679 ctgctgcacatgcctattgacatcccttatcaatgtctacagctgctggagagctgg 738
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
842 CTGCTGCACATCGTATGACATCCCTATCAATGTCTACAAGCTGCTGGCAGAGGACTGG 783

QY 739 ccatttggagctgagatgtaagctgggtcccttccatcacagaaaacctccgtggaaac 798
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
782 CCATTTGGAGCTGAGATGTATAAGCTGGTGCCTTTTCATACAGAAAGCCTCCGTGGGAATC 723

QY 799 actgtgctgagctatgtctctgagctgattgacagatatcagactgttcttggagt 858
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
722 ACTGTGCTCAGCTATGCTGCTGAGTATGACAGATATCGAGCTGTGCTTCTTGGAGT 663

QY 859 agaattaaagggaattgggttccaaaattggacagcagtagaataattgtttgattgggtg 918
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
662 AGAATTAAAGGAATTGGGGTTCACAAATGGACAGCAGTAGAATTTCTTTGATTGGGTG 603

QY 919 gtcctgtggtctggtgctgctgagctagccataggtttgatataataacgatggactac 978
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
602 GTCTCTGTGGTCTGGCTGTCCCTGAAGCAATAGGTTTGTATATAATTACGATGGACTAC 543

QY 979 aaagaaagtattctcggaatcgtctgtcttcacccgttcagaagacagctttcatgcag 1038
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
542 AAAGGAAGTTATCTGCGAATCTGCTGTCTCATCCGTTTCAGAAAGACAGCTTTCATGCAG 483

QY 1039 tttaacaagcagcaaaagattgggtggtgttcagttcttcttctgcttgccattggcc 1098
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
482 TTTTACAGACAGCAAAAGATTGGTGGCTGTTTCAGTTCTATTTCTGCTTGCCATTGGCC 423

QY 1099 atcaatgcaattttttatatacactaatgacctgtaaatgttgaagaagaaagtggcatg 1158
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
422 ATCACTGCAATTTTATATACATAATGACCTGTGAAATGTTGAGAAAGAAAGTGGCATG 363

QY 1159 cagattgctttaaatgatcacctaaagcagacagcaggggaagtggcaaaacctcttttc 1218
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
362 CAGATTGCTTTAAATGATCACCCTAAAGCAGACAGCGGAAGTGGCCAAAACCGTCTTTTC 303

QY 1219 ctggctcttgtctttgcccctgctgggtcccccctcacctcagcaggattctgaagctc 1278
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
302 CTGGTCTTGTCTTTGGCCCTGCTGGCTTCCCTTCACCTCAGCAGGATTTCTGAAGCTC 243

QY 1279 actctttataatcagaatgatcccaatagatgtaaccttttgagctttctgttggtatg 1338
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
242 ACTCTTTTAAATCAGAAATGATCCCAATAGATGTGAACCTTTTGAGCTTTTCTGTTGATATG 183

QY 1339 gactatattggtatcaaatgcttcactgaattctcgtattaaacccaattgctctgtat 1398
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
182 GACTATATTGTTATCAACATGCTTCACATGAATTTCTGCAATTAACCAATTTGCTGTAT 123

QY 1399 ttggtgagcaaaagattcaaaactgctttaagtcatgcttattgctgctgggtggcagta 1458
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
122 TTGGTGGACAAAGANTCAAAAACCTGCTTTAAGTATGCTTATGCTGCTGGTGGCCAGTCA 63

QY 1459 ttggaagaaaacagctccttggaggaagaaagcagctgctgctaaagtcaaaagctaatgat 1518
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
62 TTTGAAGAAAACAGCTCCTTTGGAGGAAAAGCAGTCTGCTTAAAGTTCAAAGCTCAATGAT 3

QY 1519 ca 1520
Db      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 NA 1
```

```
Qy 869 gaattggggttcacaaatggacagcagtagaataattgttttgattgggtggtctctgtgg 928
|||||
Db 657 GAATTGGGGTTCACAAATGGACAGCAGTAGAAATTTGTTGATTTGGGTGGTCTCTGTGG 598
Qy 929 ttctgctgctcctgaaccatagtttttgatataattacatgactacacaaagaagt 988
|||||
Db 597 TTCCTGGCTGCTCCTGAAGCCATAGGTTTGTATATTAATACGATGAGTACAAAGGAAGT 538
Qy 989 atctcgaaatctgctgtcttcacccgttcagaagacagcttttcagcagttttacaaga 1048
|||||
Db 537 ATCTCGGAATCTGCTTGTTCATCCGTTTACAGACAGCTTTCATGCGAGTTTACAGA 478
Qy 1049 cagcaaaagattgggtggtgttcagttcttatcttcgttcgttcgcatcagcactgcat 1108
|||||
Db 477 CAGCAAAAAGATTGGTGGTGTTCAGTTTCTATTTCTGCTGTCATTTGGCCATTCAGTGCAT 418
Qy 1109 tt-ttttacactaatgacctgtgaaatgttgagaagaaaagtggtcagcagattgct 1167
|||||
Db 417 TTGTTTTATACACTAATGACCTGTGAATGTTGCAAAAGAAAAGTGGCATGCGATGCT 358
Qy 1168 ttaaatgatcaactaaagcagagcgggaagtggccaaacccgtcttttgcctgctcctt 1227
|||||
Db 357 TTAATATGATCACTAAAGCAGAGACGGGAAGTGGCCAAACCGTCTTTTGGCTGCTCTT 298
Qy 1228 gctttgacctgtgcttcccttcacctcagcagagattctgaagctcactctttat 1287
|||||
Db 297 GTCTTTGGCCCTGCTGGCTTCCCTTCACCTTCAGCAGGATTCCTGAAGCTCACTCTTAT 238
Qy 1288 aatcagaatgatcccaatagatgtgaacttttgagcttctctgtgtgtgttgactatatt 1347
|||||
Db 237 AATCAGATGATCCCAATAGATGCACTTTTGACCTTCTCTGTTGATTTGACATATT 178
Qy 1348 ggtatcaacatggcttccactgaattcctgcattaccccaattgctctgtatttggtagc 1407
|||||
Db 177 GGTATCAACATGGCTTCACTGAATTCCTGCATTAACCCAAATGCTCTGTATTGCTGAGC 118
Qy 1408 aaaaatcacaacagctttaaagtcagcttatgct-gctggtgccagtcatttgaaga 1466
|||||
Db 117 AAAAGATTCAACACTGCTTTTAAGTCATGCTTATGCTGGCTGGTGGTCCAGTCAATTGAAGA 58
Qy 1467 aaaaacagtccttggaggaaaagcagtcgtcttaagtcaaaagcctaag 1516
|||||
Db 57 AAAACAGTCTTGGAGGAAAAGCAGCTGCTGCTTAAGTTTACAGCTAAGG 8

RESULT 3
LOCUS AL543805 874 bp mRNA linear EST 16-FEB-2001
DEFINITION AL543805 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI005YG20 5
prime, mRNA sequence.
ACCESSION AL543805
VERSION AL543805.1 GI:12876284
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 874)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
CONTACT Genoscope
COMMENT Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
source
location/Qualifiers
1..874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI005YG20"
/tissue_type="placenta"
```

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 298 a 161 c 139 g 270 t 6 others
ORIGIN

Query Match 19.6%; Score 842.6; DB 9; Length 874;
Best Local Similarity 98.7%; Pred. No. 7.2e-134;
Matches 864; Conservative 4; Mismatches 5; Indels 2; Gaps 2;

Qy 1821 acatagctcttaacaactccaggatattcacacaacacttaggttataaatgagctca 1880
|||||
Db 1 ACATAGCTCTTAACAACCTCCAGGATATTACACAAACACTTAGGCTTAAAAATGAGCTCA 60
Qy 1881 ctcaagaattctattcttctctaaagaagatttttttaaatcaatggagctctgatat 1940
|||||
Db 61 CTCAGAAATTTCTATTCTTTCTTCTAAAAGAGATTTATTTTAAATCAATGGGACTCTGATAT 120
Qy 1941 aaaggagaataaagtcactgtaaaacagaaacttttaaatgaagcttaaatctcaatt 2000
|||||
Db 121 AAAGCAAGAATTAAGTCACTGTAAACAGAACTTTTAAATGAAGCTTAAATTTACTCAATTT 180
Qy 2001 aaatttttaaaactcctttaaacc-aacttttcaataataattatcacactatcagat 2059
|||||
Db 181 AAAATTTTAAATCCTTTAAACTAACTTTTCAATTAATATTATATCACACTATTATCAGAT 240
Qy 2060 tgaattagatgcaaatgagagcagtttagttgttcatttttcgacactgaaaca 2119
|||||
Db 241 TGTAAATTAGATGCAAAATGAGAGAGCAGTTTAGTTGTTCATTTTCGGACACTGGAAACA 300
Qy 2120 ttttaaatgatcaggagggttaacagaaagcagagctgtttttgaaatcattcacact 2179
|||||
Db 301 TTTAAATGATCAGGAGGAGTAAACAGAAAGCAAGGCTGTTTTTGAATCATTTACACT 360
Qy 2180 ttcactagaagcccaaacctcagcattctgcaattgttaaccacacatgtcacaacaagc 2239
|||||
Db 361 TTCACTAGAAGCCCAAACTTCAGCAATTCGCAATATGTAACCAACATGTCTCACAACAAGC 420
Qy 2240 agcatgtaacagactggcacatgtgccagtcgaatttttaaaataataactttttaaaga 2299
|||||
Db 421 AGCATGTAAACAGACTGGCAGCATGTCCAGCTGAATTTTAAATATATATTAATTTTAAAGA 480
Qy 2300 aaattattacatcctttacattcagtttaagatcaaacctcacaagagagaaatagaatgt 2359
|||||
Db 481 RAATTATTACATCCTTTACATTCAGTTAAGATCAAACTCACAAGAGAGAAATAGAAATGT 540
Qy 2360 tgaagggtatcccaaaaagactttttgaatgtgctcattcacatcacctcgtggaagacaat 2419
|||||
Db 541 TGANAGGTATCCCCAAAAGACTTTTGTGAATCTGCTCNTTCTCATACCTGTGGAACAAAT 600
Qy 2420 actatctacaatttttcagattattaaaactctcttttttcacatctgtagcttaaac 2479
|||||
Db 601 ACTATCTACAAATTTTTCAGGATTTAAMATCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 660
Qy 2480 tctgtttgtgtttgtctatctgtaataacttacctacacacactgcatgtagatgattaaa 2539
|||||
Db 661 TCTGTTGGTTTGTGTCATCTGTAATACTTACCTACATACACTGTCATGTAGATGATTAAAT 720
Qy 2540 tgaggggcagccctgtgtctcatagctttacgatggagagatgccagtgacctcacaataa 2599
|||||
Db 721 TGAGGGCAGGCCCTGTGCTCATAGCTTTACGATGGAGAGATGCCAGTACCTCATCAATAAA 780
Qy 2600 agactgtgaactgctggtgcagtgctccacatgacaagggggcaggtagaccctctctc 2659
|||||
Db 781 AGACTGTGAACCTGGCTGGTGAGTGTCACATGACAAAGGGGGCAGGTAGCACCTCTCTC 840

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive, Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

BASE COUNT 207 a 251 c 273 g 239 t 2 others
ORIGIN

Query Match 18.6%; Score 798.6; DB 9; Length 972;

Best Local Similarity 99.0%; Pred. No. 2.2e-126;

Matches 823; Conservative 1; Mismatches 5; Indels 2; Gaps 2;

```

Qy 178 tgaacttggtctgaaactcgagagcgccaccggagcctctctggagcaggtaggcgc 237
|||
Db 144 TGTCTCTAGGCTCTGAACCTGGGAGCGGCCACCGGACGCTTCTGGAGCAGGTAGCAGC 203
|||
Qy 238 atgagcgctccaagtctgtgcgagcgccctggttgcgtggttcttgcctgcgc 297
|||
Db 204 ATGCAGCGGCTCCAACTCTGTGGGACGCGCCCTGGTGGCTGTCTTCCCTGGCGC 263
|||
Qy 298 ctgctcggtatctgggagagagagaggtctccgcctgacagggccactccgctttg 357
|||
Db 264 CTGTCGCGGATCTGGGAGAGGAGAGAGGCTTCCCGCTGACAGGGCCACTCCGCTTTG 323
|||
Qy 358 caaacgcagagataatgaagcaccaccactaagacctatggcccaagggttccaagcc 417
|||
Db 324 CAACACCAGAGATAATGAGCGCCACCCACTAAGACCTTATGGCCCAAGGTTTCCAAGCC 383
|||
Qy 418 agtctgcggtctgttgcgacctgcgaggtgcttaagagagacagcgaggtatct 477
|||
Db 384 AGTCTGCGGGTGTGTGGACCTGCGAGGTGCTTAAGAGACAGGACGCGCAGATCT 443
|||
Qy 478 ccgcacgcaccatctccctcccccgtgccaaaggaccatcgagatcaaggagacttc 537
|||
Db 444 CCGCCAGCACCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTC 503
|||
Qy 538 aaatacatcaacacggttgtctgctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 597
|||
Db 504 AAATACATCAACACGGTGTGTCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 563
|||
Qy 598 acactctgagaattatctcaagaacaagtgcgtgcgaacggtcccaatatctgtatc 657
|||
Db 564 ACACCTCTGAGAATATCTCAAGAACAAGTGCATGCGAAACGGTCCCAATATCTTGATC 623
|||
Qy 658 gccagcttggtctgggagacctgctgcacatcgctcatgtgacatccctataatgtctac 717
|||
Db 624 GCCAGCTTGGCTCTGGGAGACCTGCTGCACATGCTCATTTGACATCCCTATCAATGTCTAC 683
|||
Qy 718 aagctgtgcagaggactggccatttggagctgagatgtgtgaagctgggtcccttccata 777
|||
Db 684 AAGCTGTGCGAGAGACTGGCCATTTGGAGCTGAGATGTGTAAAGCTGGTGCCTTTGATA 743
|||
Qy 778 cagaaagcctctgtggaatcacgtgtgctgagctatgtctgtgtgtgtgtgtgtgtgt 837
|||
Db 744 CAGAAAGCCCTCGTGGGAATACATGTGCTGAGTGTATGTGCTGTGAGTATTGACAGATAT 803
|||
Qy 838 cgagctgtgtctcttgagtagaattaaagggaattgggttcccaaatggacagcagta 897
|||
Db 804 CGACCTGTGCTTCTTGGAGTAGAATTAAGGAATTTGGGTTT -CAAAATGGACAGCAGTA 862
|||
Qy 898 gaaattgtttgatttgggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 957
|||
Db 863 GAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 922
|||
Qy 958 gataataattacgactggaactcaaaaggagttatctgtggaattctgtgtgtgtgt 1008
|||
Db 923 GATATAATTACGATGGACTACAAAGG -AGTTATCTGCGAATCTGCTTGCCT 972
|||

```

RESULT 6

AL553041/c

LOCUS

DEFINITION

prime, mRNA sequence.

AL553041

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES

source

1. 823

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0D1072YK22"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and

cloned into the Not I and Eco RV sites of the pCMVSPORT 6

vector. Library was normalized. Library was constructed by

Life Technologies. Contact : Feng Liang Life Technologies,

a division of Invitrogen 9800 Medical Center Drive,

Rockville, Maryland 20850, USA Fax : (1) 301 610 8371

Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com"

BASE COUNT 305 a 93 c 129 g 290 t 6 others

ORIGIN

Query Match 17.5%; Score 750.6; DB 9; Length 823;

Best Local Similarity 97.3%; Pred. No. 3.5e-118;

Matches 802; Conservative 5; Mismatches 7; Indels 10; Gaps 4;

Qy 3383 tcccaatcaccttttcagagcctgttatcatagagctatttagactctcaatttaa 3442

Db 823 TCCCAATCACCTTTTTCAGAGCKCTGTATCATAGAAGTCATTTTAGACTCTCAATTTAA 764

Qy 3443 att-aattttgaatcacataattttcacagtttatlaatatatttatttctatttaa 3501

Db 763 ATTAATTTGAATCAGTAAATTTTTCACAGTTTATTAATATATTAATTTCTATTAAA 704

Qy 3502 ttttagattatttttatacatgtactgaatttttaacatcctgatacccttctcttc 3561

Db 703 TTTTAGATTATTTTATTACCATGTACTGAATTTTACATCTCTGATACCCCTTCTCTC 644

Qy 3562 catgtcagatcatgttctctctaatattcttgcacaaatttgcacacacacacacacac 3621

Db 643 CATCTCAGTATCATGTTCTCTTAATATCTTTCGCCAAATTTTGAACACTACACAAAAGCA 584

Qy 3622 tactgtcattatttataataaaatgtcattcagtggtctttttaaataaaatgttgatc 3681

Db 583 TACTTGGATATTTTATATAAATCGCATTCAGTGGCTTTT -AAAAAATGTTTGTATC 525

Qy 3682 aaaacttaacatactgataagtaagaacaattataattctttacatactcaaaacca 3741

Db 524 AAAACTTTVACATACTGATAAGTAAGAAACAATTAATAATTTCTTNACATACTCAAAACCA 465

Qy 3742 agatagaaaaaggtgctatcgcttcaactcaaacatgttcttcctagatttaaggacttta 3801

|||||

```

Db 464 AGATAGAAAGGGTCTATCGTTCAACTTCAAAACATGTTTCCTAGTATTTAAGACATTTA 405
QY 3802 atagcaacagacaaatattgttaacatggatgttacagctcaaaagatttataaaa 3861
Db 404 ATATGCMACAGACAAAATTTATGTTAAACATGGATGTTACAGCTCAAAAGATTTATAAAA 345
QY 3862 gattttaaccattttccctctattatccactgctgaatggatgtatgttcaaacacc 3921
Db 344 GATTTAAACCTATTTTCTCCCTATTATTCACCTGCTAATGTTGGATGATGTTCAACACC 285
QY 3922 ttttagtattgagcttacatattgccaaggaatacacagtttatagcaaaacatgggta 3981
Db 284 TTTTAGTATTGATAGCTTATATGCTATGCGCAAGGAATACAGTTTATAGCAAAACATGGGTA 225
QY 3982 tgcgtagctaacatttataaaagttaataaacaatgtaaaataatatatctggag 4041
Db 224 TGCTGTAGCTAACTTTATATAAGTGTAATATACAAATGTAATAAATTTATATCTGGGAG 165
QY 4042 gatttttggctgctaaagtggtctatagttactgattttttatttatgtgaagcaaaacaa 4101
Db 164 GATTTTGGTGGCTAAAGTGGCTATAGTTACTGATTTTATTATGTAAGCAAAACCA 105
QY 4102 ataaaatttaagtttttttcaaaactacatttttttttttttttttttttttttttttt 4160
Db 104 ATAAAATTTAAGTTTTCACAACTACCTTTATTTTTCACGTGTACAGACACTAATTCA 45
QY 4161 ttaataactaaa-----tgattgttttaaaagaaataataaag 4197
Db 44 TTAATACTCACACTCTCGACCTGTTTAAAGAAATATAAAG 1

RESULT 7
AL553065 796 bp mRNA linear EST 16-FEB-2001
LOCUS AL553065 LTI_NFL006_PL2 Homo sapiens cDNA clone CS0DI072YK22 5
DEFINITION prime, mRNA sequence.
ACCESSION AL553065
VERSION AL553065.1 GI:12892551
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..796
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DI072YK22"
/clone_lib="LTI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 251 a 137 c 141 g 265 t 2 others
ORIGIN

Query Match 17.4%; Score 746.6; DB 9; Length 796;

```

```

Best Local Similarity 99.1%; Pred. No. 1.7e-117;
Matches 790; Conservative 2; Mismatches 1; Indels 4; Gaps 4;
QY 2680 ggtttctagcatatgataatgctataagcttaaaatactatttttcaaaatcatcacagatt 2739
Db 1 GGTTTCTAGCATATGTATAATGCTATAAGTATAAAATACTATTATTTCAAAATCATACAGATT 60
QY 2740 agtcatattacagctacatctgaaagcttataactaaatattttgtattatttttgaataa 2799
Db 61 AGTCAATTTAACAGCTACCTGTAAAGCTTATTACAAATTTTGTATTATTTTGTAAATA 120
QY 2800 gccaatgaaagatttgccttgacatggcttttcttctcatctagagcaaaactgcttt 2859
Db 121 GCCAATGAAAGTTTGGCTTGACATGGTGCCTTTCTTCATCTAGAGCAAAACTGCTTT 180
QY 2860 ttgagacgttaagaacctcttagctttgtgcgttccctgctaatttttatctcttaag 2919
Db 181 TTGAGACCGTAAGAACCTCTTAGCTTTTGGCTTCTGCTAAATTTTATATCTTCTAAG 240
QY 2920 caaagtctcttagatagcttggatgagatgtgtgaaagatgtacaagag-aaaac 2978
Db 241 CAAAGTGCCTTAGGATAGCTTGGGATGAGATGTGTGAAAGATATGTACAAGAGCAAAAC 300
QY 2979 ggaagagagagaaatgaggtgggttgaggagaaacccatggggacagatttcccatctt 3038
Db 301 GGAAGAGAGAGAAATGAGGTGGGTGGAGGAAACCCATGGGGACAGATCCCATTCCT 360
QY 3039 agcctaagcttcgtcattgcctctgcacatcaatgcaaaagctcgtgatttggctccagc 3098
Db 361 AGCCTAAAGTTCGTCAATTCCTCGTCAACCAATGCAAAAGCTCCTGATTTTGTTCACG 420
QY 3099 aaacacagctgcaatgttctcagagtgactttcgaaataaaattgggccaagagctttaa 3158
Db 421 AAACACAGTGCATATGTTCTCAGAGTACTTTCGAAATAAATTTGGGCCCCAGAGCTTTAA 480
QY 3159 ctcggtcttaaaatagcccaatttttactttgttttcttcttaataaggctggggccaca 3218
Db 481 CTCGGTCTTAAATATGCCCAATTTTACTTTGTGTTTTCTTTTAAATAGGCTGGGCCACA 540
QY 3219 tgttgaaataagcagtagtaattgtgtttctctgcaataattgaaatgtgatacagtaaa 3278
Db 541 TGTTGGAATAAAGCTAGTAATATGTTGTTCTCTCAATATTGAATGTGATGTGTACAGTAAA 600
QY 3279 ccaaaacccaacatgtggccagaaagaaagcaataataataa-ttcacacacata 3337
Db 601 CCAAAACCCAACATGTGGCCAGAAAGAGCAATAATAATAATTTCACACACCATA 660
QY 3338 tggattctattataaatcacccac-aaaactgttctttaaattttcatcccaatcatttt 3396
Db 661 TGGATTCTATTATAAATACCCACAAACTTGTCTTTAATTTTCATCCCAATCACTTTT 720
QY 3397 tcagaggcctgttatcatagaagtcatttttagactctcaattttaataatttttgaatc 3456
Db 721 TCAGAGGYCTGTATCATAGAAAGTCATTTTAGACTCTCAATTTTAAATWAA-TTTGAATC 779
QY 3457 actaataatttcacagt 3473
Db 780 ACTAATATTTTCAGT 796

RESULT 8
LOCUS BM014035
DEFINITION 603639686F1 NIH_MGC_87 Homo sapiens cDNA clone IMAGE:5415716 5',
mRNA sequence.
ACCESSION BM014035
VERSION BM014035.1 GI:16528389
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 770)

```


Db 121 GCCTTGTTGCTGCTGGGATCATCGGAACTCCACACTTCTGAGAAATATCTACAGA 180
QY 623 acaagtgcagcaaacgctcccaatatcttgatcgccagcttgctgctggagacctgc 682
Db 181 ACAAGTGCATGCGAAACGGTCCCAATATCTGATCGCCAGCTTGGCTCTGGAGACCTGC 240
QY 683 tgcacatcgtcattgacatccctcatcaatctctacaagctgctggcagagactggccat 742
Db 241 TGCACATCGTCATGACATCCCTATCAATGCTCTACAAGCTGCTGGCAGAGACTGGCCAT 300
QY 743 ttgagctgagatgtgtaagctggtgcttccatcacagaaagcctccctgggaaatcactg 802
Db 301 TTGAGCTGAGATGCTGAAGCTGGTGCCTTCATACAGAAAGCTCCGCTGGGAATCACTG 360
QY 803 tgcctgagctatgctcctgagctgattgacagatatcgagctgttcttcttgagtagaa 862
Db 361 TGTGAGTCTATGCTGCTGAGTATTGACAGATATCGAGCTGTGCTTCTTGAGTAGAA 420
QY 863 tta-aaggaattgggttccaaatggagacagcagtagaataattgttttgatttgggtgctc 921
Db 421 TTACAGGAATTTGGGGTTCAAAATGGACAGCAGTAGAATTTGTTGATTTGGGTGGTC 480
QY 922 tctggtgtctgctgctccctgaagccataggtttttgataataattacgatggactacaaa 981
Db 481 TCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 982 ggaagtattctgcgaatctgctgcttccctgcttccctgcttccctgcttccctgcttccct 1041
Db 541 GGAAGTTATCTGGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 1042 tacaagacagcaaaagatggtgctgcttccctgcttccctgcttccctgcttccctgcttccct 1101
Db 601 TACAAGACAGCAAAAGATCGTGGCTATTCAGTTCCTATTCCTGCTGCTGCTGCTGCTGCTGCT 660
QY 1102 actgcattttttatatacaataatgacctgaaatgtaagtagaaga-aaagtggcatgca 1160
Db 661 ACTGCATTTTATACACTAATGACCTGTGAATGTTGGAAGAGCAAAAGTGGCATGCA 720
QY 1161 gattgctttaatgatcacctaaagcagagacgaggaag 1198
Db 721 GATTGCTTTACATGATCACCCTAAAGCAGAGACGGGAAG 758

RESULT 10
BG769122
LOCUS
DEFINITION BG769122 885 bp mRNA linear EST 15-MAY-2001
602743382F1 NIH_MGC_49 Homo sapiens CDNA clone IMAGE:4872909 5',
BG769122
ACCESSION BG769122.1 GI:14079775
VERSION EST.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-re@mail.nih.gov
Tissue Procurement: ATCC/DCTP/DTP
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: L1CM1750 row: c column: 22
High quality sequence stop: 816.
Location/Qualifiers
I..885
/organism="Homo sapiens"

/db_xref="taxon:9606"
/clone="IMAGE:4872909"
/clone_lib="NIH_MGC_49"
/tissue_type="melanotic melanoma, high MDR (cell line)"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pONB7; Site_1: XhoI; Site_2:
ECORI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGCAG(G). Size-selected >500bp for average insert size
1.8kb. Library constructed by Ling Hong in the laboratory
of Gerald M. Rubin (University of California, Berkeley)
using ZAP-cDNA synthesis kit (Stratagene) and Superscript
II RT (Life Technologies). Note: this is a NIH_MGC
Library. I"
BASE COUNT 213 a 211 c 225 g 236 t
ORIGIN
Query Match 16.9%; Score 727.8; DB 10; Length 885;
Best Local Similarity 95.1%; Pred. No. 2.7e-114;
Matches 838; Conservative 0; Mismatches 32; Indels 11; Gaps 8;
QY 383 ccactaagaccttatgcccagggttccaaagcagctctgcgcgctcgttggcacctg 442
Db 1 CCACCTAAGACCTTATGCCCCAAGGGTTCCAAACGCCAGTCTGGCGGGTGGTGGACCTG 60
QY 443 cggaggtgcctaaaggagacagcgagcagctatcccgccacgacacctctccctcccc 502
Db 61 CGGAGGTGCTTAAGGAGACAGGAGCGGAGGATCTCCGCCACGACCATCTCCCTCCCC 120
QY 503 cgtgccaagaccatcgagatcaagagagactttcaaatatatacaacacggttggctct 562
Db 121 CGTGCCAAAGACCCATCGAGATCAAGGAGACTTTCAAAATACATCAACACGGTGTGTCT 180
QY 563 gccttggtctgctgctgggagatcatcggaactccacactctctgagaattatctacaaga 622
Db 181 GCCTTGTGTCGTGCTGGGATCATCGGAACTCCACACTTCTGAGAAATATCTACAAGA 240
QY 623 acaagtgcagcaaacgctcccaatatcttgatcgccagcttgctcctggagacctgc 682
Db 241 ACAAGTGCATGCGAAACGGTCCCAATATCTGATCGCCAGCTTGGCTCTGGGAGACCTGC 300
QY 683 tgcacatcgtcattgacatccctcatcaatgtctacaagctgctggcagagactggccat 742
Db 301 TGCACATCGTCATGACATCCCTATCAATGCTCTACAAGCTGCTGGCAGAGACTGGCCAT 360
QY 743 ttgagctgagatgtgtaagctggtgcttccatcacagaaagcctccgctgggaatcactg 802
Db 361 TTGAGCTGAGATGCTGTAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
QY 803 tgcctgagctatgctcctgagctgattgacagatatcgagctgttcttcttgagtagaa 862
Db 421 TGTGAGTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 863 ttaaggaattgggttcccaaatggacagcagtagaataattgttttgatttgggtgctc 922
Db 481 TTAAGGAACCTGGGTTTCCAAAATGGACAGCAGTAGAAA-TGTGCTGAATTTGGTGGTCT 539
QY 923 ctgtggtctcgtgctccctgaagccataggttttctgataataattacagtagactacaag 982
Db 540 CTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 599
QY 983 gaagtatctgcgaatctgcttcttccctgcttccctgcttccctgcttccctgcttccctgctt 1041
Db 600 GAAGTTATCTGCGAAATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 659
QY 1042 tacaagacagcaaaagattggtgctgcttcttcttcttcttcttcttcttcttcttcttct 1099
Db 660 TACAAGACAGCAACAGA-TGGTGGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 718
QY 1100 tcaactg--cattttttatatacaactaatgacc--tgtgaattgttgagaag--aaaagt 1153
Db 719 TCACTGGCATTTTTTTTATACACTAATGACCTGTGAACATGTTGAGAAAGGAAAGTGG 778

Db	602	TATTCTGCTGGCATTGCGCATCTACACCTAATATGACCTGTGAATG	543
Qy	1138	ttgagaagaaaagtgcctgcagattgcttttaaatg--atcacctaaagcagagacggg	1195
Db	542	TTGAGAAAGAAAGTGCATGTCAGATGCTTTAAATNGATNCACCTAAAGCAGAGACGGG	483
Qy	1196	aagtggccaaaacgg--tcctttgctgggtccttctgtcttctgtcctctgctggctccct	1253
Db	482	AAGTGGCCAAACCGCTGCTTTGGCTGGCTCTGCTTTGGCCCTCTGCTGGCTCCCT	423
Qy	1254	tcacctcagcagattcctgaagctcacctcttataatcagaatgattcccaatagatga	1313
Db	422	TCACCTCAGCAGGATTCGAAAGCTCACTCTTTAATCAGAATGATCCCAATAGATGTG	363
Qy	1314	actttgagctttctggt--ggtattgactatattggtatcaacatgcttc--actgaat	1371
Db	362	ACTTTTGAGCTTTCTGTTTCGGTATTGGACTATATTGGTATCAACATGGCTTCAACTGA	303
Qy	1372	tcctgcattaccacaattgctctgatttggtagcagaaagattcaaaaactgctttaag	1431
Db	302	GCKGKCATTAACCCAAATGGCTCTGCTATTGGTGAGCAAAAGATTCAAAAAC	243
Qy	1432	tcctgcttatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctg	1491
Db	242	TCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	183
Qy	1492	tcgtgcttaagttcctgaagctaatgacaggtatgacacattccgttcctcagtaataa	1551
Db	182	TCGTGCTTAAAGTTCAAAGCTAATGATCACGGATATGACCACTTCGCTTCCAGTAATAA	123
Qy	1552	taagctcatcttgaaagagaactatcactgctatttccttcttcttcttcttcttcttctt	1611
Db	122	TACAGCTCATCTTGAAAGAGAGACTATTCACGTGATTTCATTTCCTTTATATTGGACGA	63
Qy	1612	agtcattaaaacaaaatgaaacatttgcacacacacacacacacacacacacacacacac	1663
Db	62	AGTCATTAACAAATGANAATTTGCCAAACAAACAAACAAACAAACAAACAAACAAAC	11
RESULT	12		
LOCUS	AU117045		
DEFINITION	AU117045 HEMBA1 Homo sapiens cDNA clone HEMBA1000541 5', mRNA	785 bp	linear EST 19-OCT-2000
ACCESSION	AU117045		
VERSION	AU117045.1		
KEYWORDS	EST		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.		
TITLE	HRI human cDNA project		
JOURNAL	Unpublished (2000)		
COMMENT	Contact: Takao Isogai Genomics Laboratory Helix Research Institute 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan Tel: 81-438-52-3951 Fax: 81-438-52-3952 Email: genomics@hri.co.jp HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.		
FEATURES	Location/Qualifiers		
SOURCE	1..785		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="HEMBA1000541"		

598 CGGAGACTCCACACTTCTGAGAAATTATCTACAGAAACAAGTCATGCGAA 637
117 eGlyAsnSerThrLeuLeuArgIleIleTyrlYsAsnLysCysMetArg 134
638 ACGTCCCAATATCTGATCCCAAGCTTGCTCTGGGAGACCTGCTGCAC 687
134 sNgiPyroAsnIleuValAlaSerLeuAlaLeuGlyAspLeuLeuHis 150
688 ATCTGCTATTGACATCCCTATCAATGTCTACAAGCTGCTGCAGAGACTG 737
151 IleValIleAspIleProIleAsnValTyrlYsLeuLeuAlaGluAspTr 167
738 GCCATTGGAGCTGAGATGTGAAGCTGGTGGCTTCATACAGAAAGCCT 787
167 pPropheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGTGGGAATCACCTGCTGAGTCTATGTCTCTGACTATTGACAGATAT 837
184 eValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrl 200
838 CGAGCTGTGCTCTCTGGAGTAGAATTAAGAGGAATGGGGTTCACAAATG 887
201 ArgAlaValAlaSerTrpSerArgIleLysGlyValProLysTr 217
888 GACAGACTAGAAATCTTTGATTGGTGGTCTCTGTGGTCTTGCGTG 937
217 pThrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVal 234
938 TCCCTGAAGCCATAGGTTTTCATATATATTACGATGGACTACAAAGGAAGT 987
234 alProGluAlaIleGlyPheAspIleIleThrMetAspTyrlYsGlySer 250
988 TATCTGCCAATCTGCTGCTTCATCCGTTTCAGAAAGACAGCTTTCATGCA 1037
251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetG 267
1038 GTTTTACAAGACAGCAAAAGATTGGTGGCTGTTCCAGTTTCTATTCTGCT 1087
267 nPheTyrlYsThrAlaLysAspTrpTrpLeuPheSerPheTyrlPheCysL 284
1088 TGCCATTGGCCATCACATCCATTTTTTATACACTAAATGACCTGTGAAATG 1137
284 euProLeuAlaIleThrAlaPhePheTyrlThrLeuMetThrCysGluMet 300
1138 TTGAGAAAGAAAAGTGGCATCCAGATTGCTTTAAATGATCACCTFAAAGCA 1187
301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
1188 GAGACGGGAAGTGGCCAAAACCGTCTTTTGCTGGTCCCTGTGCTTTGCC 1237
317 nArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaL 334
1238 TCTGCTGCTTCCCTTCACTCAGCAGGATTCTGAAGCTCAGCTCTTTAT 1287
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyrl 350
1288 ATACAGATTGATCCCAATAGATGTGAACCTTTTGAGCTTTCTGTGGTGATT 1337
351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuValLe 367
1338 GGACTATATTGGTATCAACATGGGTTCACTGAATTCCTGCAATTAACCCAA 1387
367 uAspTyrlIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnProI 384
1388 TTGCTCTGTATTTGGTGAGCAAAAGATTCAAAAACCTGCTTTAAGTCATGC 1437
384 LeAlaLeuTyrlLeuValSerLysArgPheLysAsnCysPheLysSerCys 400
1438 TTATCTCTGCTGGTGCACATCTTTCAAGAAAAACAGCTCTTGGAGGAAA 1487
401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGluL 417
1488 GCAGTCGTGCTTAAAGTTTCAAAGCTTAATGATCACGGATATACAACTCC 1537

```

|||||
417 sGlnSerCysLeuPheLysAlaAsnAspHisGlyTyrAspAsnPheA 434
1538 GTTCCAGCTAATAATACAGCTCATCT 1563
|||||
434 IgSerSerAsnLysTyrSerSerSer 442
|||||

seq_name: sp_mammal:Q9N0W7

seq_documentation_block:
ID Q9N0W7 PRELIMINARY; PRT; 441 AA.
AC Q9N0W7;
DT 01-OCT-2000 (TReMBLrel..15, Created)
DT 01-OCT-2000 (TReMBLrel..15, Last sequence update)
DT 01-JUN-2001 (TReMBLrel..17, Last annotation update)
DE ENDOPHELIN B RECEPTOR.
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OC NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A.
RL Yang L., Wu Q., Yang N., Long Q., Wang X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF245469; AAF60366.1; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECIP_F1_1; UNKNOWN_1.
DR PROSITE; PS02662; G_PROTEIN_RECIP_F1_2; 1.
DR Receptor.
KW SEQUENCE
SQ SEQUENCE 441 AA; 49441 MW; 059BF28017E02A19 CRC64;

alignment_scores:
Quality: 2113.50 Length: 443
Ratio: 4.996 Gaps: 2
Percent Similarity: 95.485 Percent Identity: 90.293

alignment_block:
US-09-931-157-2 x Q9N0W7

Align seg 1/1 to: Q9N0W7 from: 1 to: 441

238 ATGCAGCGCGCTCCAAAGTCTGTGGAGCGCGCTGTTCGCTGTTCT 287
|||||
1 MetGlnProProSerLeuCysGlyLeuAlaLeuAlaLeuVal 17
|||||

288 TGCCTGCGGCGTGTGCGGATCTGTGGAGAGAGAGAGCTCCCGCTG 337
|||||
17 uAlaCysGlyMetAlaGluValTrrpGlyGluGluArgGluMetProSerA 34
|||||

338 ACAGGGCCACT...CCGCTTTTCCAAACCCGACAGATATACGCCACCC 384
|||||
34 laProAlaThrProProLeuLeuGlyAlaSerGluLeuLeuThrProSer 50
|||||
385 ACTAAGACCTTATGGCCCAAGGGTTCACCGCCAGTCTGCGCGCTGCT 434
|||||
51 ThrLysThrSerTrrpProArgspSerAsnAlaSerLeuProArgSerSe 67
|||||

435 GGCACCTCGCGAGGTGCTTAAGAGACAGAGCGGAGGATCTCCGCCAC 484
|||||
67 rAlaProAlaGluIleProLysGluGlyArgThrAlaGlyAlaProArgA 84
|||||

485 GCACCATCTCCCTCCCGCTGCCAAGGCCCATCCAGATCAAGAGACT 534
|||||
84 rgThr.....ProProCysGlnArgProThrGluIleLysAspThr 98
|||||

535 TTCAAAATACATCAACACGGTGTGTCTGCTGTGTCTGCTGCTGGGAT 584
|||||
99 PheLysTyrIleAsnThrValSerCysLeuValPheValLeuGlyI 115
|||||

585 CATCGGGAACCTCCACACTCTTGAGATTATCTACAAGAACAAAGTCATGC 634
|||||

```

|||||eileGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMeta 132
635 GAAACGGTCCCAATAATCTTGATGCCAGCTTGCTGCTGGAGACCTGCTG 684
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
685 CACATCGCATTCAGATCCCTATCATATGTCTACAAGCTGCTGGCAGAGGA 734
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
149 HisIleIleIleAspIleProIleAsnValTyrLysLeuLeuAlaGluAs 165
735 CTGGCCATTCGAGCTGAGATGTGAAGCTGGTGGCTTTCATACAGAAAG 784
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
165 ptrpProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysA 182
785 CCTCGCTGGGAATCACTGCTGAGTCTATGTGCTGCTGAGTATTGACAGA 834
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
182 IaSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 198
835 TATCGAGCTGTTGCTTCTTGGAGTGAATTAAGGAATGGGGTCCAAA 884
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
199 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProly 215
885 ATGGACAGCAGTAGAATGTTTTCATTTGCTGCTGCTGCTGCTGCTGCTG 934
215 strpThrAlaValGluIleValLeuIleTrpValValSerValIleLeuA 232
935 CTGTCCTCAAGCCATAGTCTTGTATATTAATACGATGACATACAAAGGA 984
232 IaValProGluAlaIleGlyPheAsnLeuValThrIleAspTyrLysGly 248
985 AGTTATCTCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
249 SerTyrLeuArgIleCysLeuLeuAsnProThrGlnLysThrAlaPheMe 265
1035 GCAGTTTTCACAGACAGCAAAAGATGTTGGTGTGCTGCTGCTGCTGCTG 1084
265 tGlnPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheC 282
1085 GCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
282 yLeuProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGlu 298
1135 ATGTTGAGAAAGAAAGTGGCATGAGATGCTTTAAATGATCACCATAA 1184
299 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 315
1185 GCAGACGGGAAGTGGCCAAACCGTCTTTTGCCTGCTGCTGCTGCTGCTG 1234
315 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuAlPheG 332
1235 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1284
332 yLeuCysTrpLeuAlaLeuAlaLysLeuSerArgIleLeuLysLeuThrLeu 348
1285 TATAATCAGAAATGATCCCAATAGATGTGAACCTTTTGGCTTTCTGTTGT 1334
349 TyrAspGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuVa 365
1335 ATTGAGCTATATTGTTATCAACATGGCTTCACTGATTCCTGCTGCTGCTG 1384
365 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 382
1385 CAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1434
382 roIleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSer 398
1435 TGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484
399 CysLeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGl 415
1485 AAAGCAGTCTGCTTAAAGTTCAAAGCTAATGATCAGCGATATGACAAC 1534
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148

415 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsn 432
1535 TCCGTTCCAGTAATAATACAGCTCATCT 1563
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
432 heArgSerSerAsnLysTyrSerSer 441
seq_name: sp_mammal:Q9MYZ5
seq_documentation_block:
ID Q9MYZ5 PRELIMINARY; PRT; 434 AA.
AC Q9MYZ5;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENDOTHELIN B RECEPTOR (FRAGMENT).
GN EDNRB.
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Metallinos D.L., Rine J.;
RT "Exclusion of EDNRB and KIT as the Basis for White Spotting in Border
RT Collies.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF276427; AAF81902.1; -;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1
DR PRINTS; PR00237; GPCR_RHODOPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 434 434
SQ SEQUENCE 434 AA; 48753 MW; 9C698DE15FF6A5D0 CRC64;
alignment_scores:
Quality: 2045.50 Length: 434
Ratio: 4.989 Gaps: 1
Percent Similarity: 94.470 Percent Identity: 89.171
alignment_block:
US-09-931-157-2 x Q9MYZ5 ..
Align seg 1/1 to: Q9MYZ5 from: 1 to: 434
238 ATGACGCGCCCTCCCAAGTCTGTGGAGCGCGCCCTGGTTCCTGCTGCT 287
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
1 MetGlnProLeuLeuSerTrpCysArgArgAlaLeuValAlaLeuIle 17
288 TCCTGCGGCGCTGCTGCGGATCTGGGAGAGGAGAGGAGGCTTCCGCGCTG 337
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
17 uAlaCysSerGlyAlaGlyValGlnGlyGluArgGlyPheProProA 34
338 ACAGCGCCACT...CCGTTTTCGCAACCGCAGAGATATGATGCGCCACC 384
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
34 IaArgAlaThrValProLeuTrpGlyProGlyGluIleThrProPro 50
385 ACTAAGACCTTATGCCCCAAGGTTCCAAACGCTGCTGCGCGCTGCTGCT 434
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
51 ThrGluAlaSerTrpSerLysGlyProAsnAlaSerValProTrpSer 67
435 GCACCTCGGAGGCTGCCCTAAAGGAGAGGAGGAGGAGGATCTCCGCCAC 484
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
67 rAlaProProGlnMetProLysGlyArgThrAlaGlyGlyProProA 84
485 GCACCATCTCCCTCCCGCTGCCAAGGAGGAGGAGGAGGATCGAGATCAAGGAGCT 534
|||||eileGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 148
84 rGlyThrProProProCysGluLysSerIleGluLysGluThr 100
535 TTCAATACATCAACACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 584

```
|||||
101 PheLysTyrIleAsnThrValValSerCysLeuValPheValLeuGlyI 117
585 CATCGGGAACCTCCACACTTCTGAGATTATCTACAAAGAACAGTCATGC 634
|||||
117 eileGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMeta 134
635 GAAAGCGTCCCAATATCTTGTATCGCCAGCTTGCTCTGGAGACCTGCTG 684
|||||
134 rGAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 150
685 CACATCGTCATTCACATCCCTATCAATGCTTACAACTGCTGGCAGGA 734
|||||
151 HisIleIleAspIleProIleThrValTyrLysLeuLeuAlaGluAs 167
735 CTGGCCATTGGAGCTGAGATGTGAAGTGGTCCCTTTTCATACAAAG 784
167 prpProPheGlyValGluMetCysLysLeuValProPheIleGlnLysA 184
785 CTCCGCTGGGAATCACTGTGCTGAGTCTATGTCTGTGATTTGACAGA 834
|||||
184 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 200
835 TATCGAGCTGCTCTTCTTGGAGTAGAATTAAGGAATTGGGTTCCAAA 884
|||||
201 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLy 217
885 ATGGACAGCAGTAGAATTTGTTGATTTGGTGGTCTCTGTGTTCTGG 934
217 strpThrAlaValGluIleValLeuIleTrpValValSerValValLeuA 234
935 CTGCTCCCTGAAGCCATAGCTTTTGTATATTAATACGACTGACTACAAGA 984
234 laValProGluAlaValGlyPheAspMetIleThrIleAspTyrLysGly 250
985 AGTTATCTCGGAATCTGCTGCTTCATCCGTTTCAGAACAGACTTTCAT 1034
251 ArgTyrLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMe 267
1035 GCAGTGTTCACAGACAGCAAGATGGTGGCTGTTCACTTCTCTATTCT 1084
267 tGlnPheTyrLysThrAlaLysAspTrpIlePheSerPheTyrPheC 284
1085 GCTGCCATTGGCCACTGCTGCTATTTTATACACTAATGACCTGTGAA 1134
284 ysLeuProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGlu 300
1135 ATGTTGAGAAAGAAAGTGCATGCGATGCGATTGCTTTAAATGATCACCTAA 1184
301 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 317
1185 GCAGACAGCGGAAGTGGCCAAACCGTCTTTGCTGCTGCTGCTTTG 1234
317 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 334
1235 CCCTCTGCTGGCTTCCCTTCACTCAGCAGGATTCCTGAAGTCACCTCT 1284
334 laLeuCystTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrIle 350
1285 TATAATCAGATGATCCCAATAGATGAACTTTTGAGCTTTCTGTGTGT 1334
351 TyrAspGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuVa 367
1335 ATTGGACTATATTGGTATCAATGCTTCACCTGCTGCTGCTGCTGCTG 1384
367 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnP 384
1385 CAATTGCTCTGATTTGGTGAGCAAGATTCAAAACCTCTTAAGTCA 1434
384 rolleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSer 400
1435 TGCTTATGCTGCTGGTGGCAGTCAATTTGAAGAAAAACAGTCCCTTGGAGGA 1484
|||||
```

```
401 CysLeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluG 417
1485 AAAGCAGTCGTCTTAAAGTTCAAAGTAAATGATCAGCGATATGACAACT 1534
|||||
417 ulysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnp 434
1535 TC 1536
||
434 he 434

seq_name: sp-vertebrate:Q9DGM2

seq_documentation_block:
ID Q9DGM2 PRELIMINARY; PRT; 426 AA.
AC Q9DGM2;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE ENDOTHELIN RECEPTOR BL.
OS Brachydanio rerio (zebrafish) (Zebra danio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;
OC Cypriniformes; Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Parichy D.M., Kelsh R.N., Mellgren E., Lopes S., Rawls J.,
RA Johnson S.L.;
RT "Mutational analysis of roles for endothelin receptor b1 during neural
crest and pigment pattern development in the zebrafish, Danio rerio."
RL Submitted (JUN-2000) to the EMBL/GenBank/DDB databases.
DR EMBL; AF275636; AAG00977.1;
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1;
DR PRINTS; PS00237; GPCRHOOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT VARIANT 156 156 C -> *.
FT VARIANT 184 184 A -> E.
FT VARIANT 318 318 W -> *.
SQ SEQUENCE 426 AA; 48770 MW; F92C0B0B3046E4A3 CRC64;
```

alignment_scores:

Quality: 1534.00 Length: 394
Ratio: 4.383 Gaps: 4

Percent Similarity: 88.832 Percent Identity: 72.589

alignment_block:

US-09-931-157-2 x Q9DGM2 ..

Align seg 1/1 to: Q9DGM2 from: 1 to: 426

```
370 ATAATGACGCCACCCACTAAGACCTTATGCCCCAAGGGTTCCACGCCAG 419
      |||||
18 LeuLeuThrGluHisIleAlaValMetSerAlaGlnGlyLysAspPheAs 34
      |||||
420 TCTGGCGCGG...TCGTTGCCACTT.....GCGGAGG 448
      |||||
34 nGlnSerArgLeuSerMetGlyProLeuSerProThrGlnLysSerThrI 51
      |||||
449 TGCTAAAGGAGACAGCAGCGCAGGATCTCCGCCACGC.....ACCATC 492
      |||||
51 leValIleGlyAsnGlnIleAsnGluSerMetProArgArgProLysVal 67
      |||||
493 TCCCTCTCCCGGTGCCAAGACCGACGATCAAGGAGACTTTCAAATA 542
      |||||
68 LeuProPrometCysThrAspProThrGluIleArgAspThrPheLysTy 84
      |||||
543 CATCAACACGGTTGTCTCTGCTGCTTGTGTCGTGCTGGGATCATCGGA 592
      |||||
84 rIleAsnThrValValSerCysLeuValPheValValGlyIleIleGlyA 101
```

```
593 ACTCCACACTTCTGTGAGATTATCTACAAGACAACTGCTGATGGCAACGGT 642
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
101 snSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArgAsnGly 117
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
643 CCCAATATCTTGATCGCCAGCTTGCTGCTGGAGACCTGCTGCACATCGT 692
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
118 ProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHisIle 134
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
693 CATTGACATCCCTATCAATGCTACAAGCTGCTGCAGAGGACTGCCCAT 742
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
134 tileAspIleProIleAsnValTyrLysLeuLeuAlaLysAspTrpProp 151
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
743 TTGGAGCTGAGATGTGAAGCTGGTGGCTTTCATACAGAAAGCCCGTG 792
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
151 heGlyValGlyLeuCysLysLeuValProPheIleGlnLysThrSerVal 167
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
793 GGAATCAGCTGCTGAGCTATGCTCTGCTGAGTATTGACAGATATCGAGC 842
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
168 GlyIleThrIleLeuSerLeuCysAlaLeuSerIleAspArgPheArgAl 184
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
843 TGTGTGCTCTTGGAGTAGAATTAAGGAATGGGTTCAAAATGGACAG 892
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
184 aValSerSerTrpAsnArgIleLysGlyIleGlyValProLysTrpThrA 201
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
893 CAGTAGAAATCTTTTGATTTGGGTGGTCTCTGTGCTTCTGCTGCTCCCT 942
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
201 laileGluIleIleLeuIleTrpValLeuSerIleIleLeuAlaValPro 217
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
943 GAAGCATAGCTTTTCATATATATACGATGGAGTCAAGAGGAATATCT 992
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
218 GluAlaIleAlaPheAspMetIleThrMetAspTyrLysGlyGluGlnLe 234
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
993 GCGAATCTGCTGCTTCATCCCGTTCAGAAAGACAGCTTTTCATGCGAGT 1042
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
234 uArgIleCysLeuLeuHisProLysGlnArgIleLysPheMetGlnPheT 251
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1043 ACAAGACACAAAGATTGGTGGCTGCTTTCAGTTCATTCTTCTGCTGGCA 1092
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
251 yrlLysLysAlaLysAspTrpTrpLeuPheSerPheTyrPheCysMetPro 267
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1093 TTGGCCATCAGTCATTTTATACACTAATGACCTGTGAATGTTGAG 1142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
268 LeuThrCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuAr 284
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1143 AAAGAAAGTGGCATGCAGATGCTTTAAATGATCACCTAAAGCAGAGAC 1192
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
284 gLysLysAsnGlyValGlnIleAlaLeuSerAspHisIleLysGlnArgA 301
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1193 GGAAGTGGCCAAACCGCTCTTTGCTGCTGCTTCTGCTTTCGCCCTCTGC 1242
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
301 rgGluValAlaLysThrValPheCysLeuValLeuValPheAlaLeuCys 317
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1243 TGGCTTCCCTTCACCTCAGCAGGATTCGAAGCTCACTCTTTATATCA 1292
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
318 TrpLeuProLeuHisLeuSerArgIleLeuGlnArgThrIleTyrAspGl 334
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1293 GAATGATCCCAATAGATGCACTTTTGAGCTTTCTGTTGGTATTGGACT 1342
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
334 uArgAspProAsnArgCysGluLeuLeuSerPhePheLeuValLeuAspT 351
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1343 ATATTGGTATCAACATGGCTTCACTGAATTCCTGCAATTAACCAATGCT 1392
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
351 yrlIleGlyIleAsnMetAlaSerValAsnSerCysIleAsnProIleAla 367
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1393 CTGTATTGGTAGCAAAAGATTCAAAACTGCTTTAAGTCATGCTTATG 1442
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
368 LeuTyrMetValSerLysArgPheLysSerCysPheArgSerCysLeuCy 384
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1443 CTGCTGGTGCAGCTCATTTGAACAAAAACAGTCTTGGAGGAAACAGT 1492
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
384 scYsTrpCysLeuPro...ProGluIleLeuAlaMetAspAspLysGlnS 400
```

```
1493 CGTGTCTAAAGTTCAAAGCTAATGATCACGGA 1524
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
400 erCysIleLysLeuLysValThrGluArgGly 410
seq_name: sp_vertibrate:Q9DEC5

seq_documentation_block:
ID_Q9DEC5 PRELIMINARY; PRT; 413 AA.
AC Q9DEC5;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)
DE ENDORHELIN RECEPTOR.
GN ENDR.
OS Oryzias latipes (Medaka fish).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
OC Belontiiformes; Adrianichthyidae; Oryziinae; Oryzias.
OX NCBI_TaxID=8090;
RN [1]
RP SEQUENCE FROM N.A.
RA Tobita-Teramoto T., Akiyama T.;
RT "Endothelin Receptor in Medaka Melanophores. I. Cloning and its.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB045356; BAB20097.1; -.
DR InterPro; IPR00276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 413 AA; 46752 MW; 92E99424CEBA854C CRC64;

alignment_scores:
Quality: 1457.00 Length: 392
Ratio: 4.235 Gaps: 4
Percent Similarity: 87.755 Percent Identity: 69.133

alignment_block:
US-09-931-157-2 x Q9DEC5 ..
Align seg 1/1 to: Q9DEC5 from: 1 to: 413
400 CCCAAGGGTTCACAGCCAGCTGCGCGGTGCTGGCCACCTGCGGAGGT 449
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
29 ProGluAlaLeuGluValSerGluThrAsnSerSerValGlnLeuProLe 45
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
450 GCCTAAAGGAGACAGAGCGGAGGATCTCCGCCACGCCACCATCTCCCTC 499
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
45 uProGlnProGlyArg.....ProArgSerProProp 57
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
500 CCCCGTGCACAGAGCCATCGAGATCAAGGAGACTTTCAATACATCAAC 549
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
57 roMetCysValLysProThrGluIleLysAspAlaPheLysTyrValAsn 73
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
550 ACGGTTGTCTCTGCTTGTCTGCTGGGATCATCGGAACCTCCAC 599
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
74 ThrIleSerCysLeuIlePheValValGlyIleIleGlyAsnSerTh 90
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
600 ACTTCTGAGAATTATCTACAAGCAAGTGCATGCGAAACGGTCCCAATA 649
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
90 rLeuLeuArgIleIleTyrArgAsnLysCysMetArgAsnGlyProAsnV 107
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
650 TCTTGATCCCGACCTGCTGCGGAGACCTGCTGCACATCGTCATTGAC 699
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
107 alIleuGlySerLeuAlaLeuGlyAspLeuLeuTyrIleIleAla 123
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
700 ATCCCTATCAATGCTTACAGCTGCGGAGACTGCGCAGAGGACTGGCCATT 749
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
124 IleProIleAsnValTyrLysLeuIleAlaGluAspTrpPheGlyVa 140
```

[illegible]

```
seq_documentation_block:
ID 073868 PRELIMINARY; PRT; 436 AA.
AC 073868;
```

679 CTGTCGACATCGTCATTGACATCCCTATCAATGTCTACAAGCTGCTGGC 728
|||||
137 LeuLeuTyrIleLeuIleAlaLeuProIleAsnValTyrLysLeuLeuAl 153
729 AGAGACTGGCCATTGGAGCTGAGATGTGAAGCTGGTGGCTTCATAC 778
153 aLysAspTrpProPheGlyValGlnValCysLysLeuValProPheIleG 170
779 AGAAGCCCTCGTGGGAATCACTGCTGAGTCTATGTGCTGCTGAGTATT 828
170 InLysAlaSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIle 186
829 GACAGATATCGAGCTGTTGCTCTCTGGAGTAGAATTAAGGAATGGGCT 878
187 AspArgTyrArgAlaValAlaSerTrpSerArgIleGlnGlyIle 203
879 TCCAAATGGACACAGCAGTAGAATTTGATTGGTGGTCTCTGCTGG 928
203 eProMetTrpLysAlaValGluValThrLeuIleTrpAlaValAlaIleV 220
929 TTCTGGCTGCTCCCTGAACCCATAGTTTGTATATAATTACGATGGACTAC 978
220 aIleuAlaValProGluAlaIleAlaPheAspMetValGluIleAsnTyr 236
979 AAAGGAATTTATCTCGGAATCTGCTGCTTCATCCCGTTCAGAACACAGC 1028
237 TrpAspGlnAspLeuTrpValCysMetLeuAlaSerGluGlnLysSerSe 253
1029 TTTCATGAGTTTACAGACAGCAAGATGCTGGCTGCTCACTTTCT 1078
253 rPheMetMetPheTyrArgAspValLysAspTrpTrpLeuPheGlyPheT 270
1079 ATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1128
270 yrPheCysLeuProLeuValCysThrGlyIlePheTyrThrLeuMetSer 286
1129 TGTGAATGTTGAGAAAGAAAGTGGCATGCAGATGCTTTAAATGATCA 1178
287 CysGluMetLeuSerLysArgAsnGlyMetArgIleAlaLeuAsnAspHi 303
1179 CCTAAGCAGACGCGGAAGTGGCCAAACCGTCTTTGCTGCTGCTGCTGCT 1228
303 SmetLysArgArgGluValAlaLysThrValPheCysLeuValValI 320
1229 TCTTTCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1278
320 lePheAlaLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLys 336
1279 ACTCTTTTATATCAGAAATGATCCCAATAGATGTGAAGTCTGCTGCTTCT 1328
337 ThrIleTyrAspGlnThrAspProAsnArgCysGluLeuLeuSerPheLe 353
1329 GTTGCTATTGGACTATTGCTATCAACATGCTTCACTGAATTCCTGCA 1378
353 uLeuValMetAspTyrPheGlyIleAsnMetAlaSerLeuAsnSerCysI 370
1379 TTAACCCCAATTCCTGCTATTGCTGAGCAAAAGATTCAAAACATGCTTT 1428
370 leAsnProValAlaLeuTyrPheValSerArgLysPheLysAsnCysPhe 386
1429 AAGTCATGCTTATGCTGCTGGGCGCAGTATTTGAAGAAACACATCCTTT 1478
387 GlnSerCysLeuCysCysTrpCysGln.....ArgProAlaLe 399
1479 G.....GAGGAAAAACAGTCGCTGCTTAAAGTCAAGCTA 1513
399 uSerIleThrProThrAspGluLysGlySerValGlyLysTrpLysAlaA 416
1514 AT.....GATCAGGATATGAC.....AACTTCGCTGCCAGTATATAA 1551
416 snGlyGlnGluLeuGlyLeuAspArgSerSerArgLeuSerAsnLys 432

1552 TACAGCTCATCT 1563
|||||
433 TyrSerSer 436
seq_name: sp_vertibrate:073739
seq_documentation_block:
ID O73739 PRELIMINARY; PRT; 421 AA.
AC O73739;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ENDOTHELIN TYPE A RECEPTOR.
GN EDNR.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=WHITE LEHORN;
RX MEDLINE=99030274; PubMed=9811577;
RA Kempf H., Linares C., Corvol P., Gasc J.M.;
RT "Pharmacological inactivation of the endothelin type A receptor in the
RT early chick embryo: a model of mispatterning of the branchial arch
RT derivatives";
RL Development 125:4931-4941(1998).
DR EMBL: AF040634; AAC7793.1;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm.1; 1
DR PRINTS: PR00237; GPCRHOOPS.
DR PROSITE: PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE: PS00262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 421 AA; 47713 MW; 33A2C97A6DCD02DF CRC64;

alignment_scores:
Quality: 1261.00 Length: 393
Ratio: 3.928 Gaps: 7
Percent Similarity: 81.679 Percent Identity: 62.341
alignment_block:
US-09-931-157-2 x O73739 ..
Align seg 1/1 to: O73739 from: 1 to: 421
409 TCCACAGCCAGTCTGCGCGGTGGTGGCACCTGCGGAGGTGCTAAAGG 458
39 ThrSerSerGlyIleGluSerSerLeuLeuPro..... 49
459 AGACAGGACGCGAGGATCTCCGCCACGACCATCTCCCTCCCGTGGC 508
50ThrThrArgArgProProValSerAsnGlnThrValLysCys 64
509 AAGGACCCATCCAGATCAAGGAGACTTCAAAATACATCAACACGTTTGTG 558
64 erGlnGlnThrLysIleAlaGluThrPheLysTyrIleAsnThrValVal 80
559 TCCTCCCTTGTGTTGCTGCTGGGATCATCGGAACTCCACACTTCTGAG 608
81 SerCysAlaIlePheIleValGlyMetValGlyAsnAlaThrLeuLeuAr 97
609 AATTATCTACAAGAACAGTGTGTCGAAACGTCCTCAATATCTTGTATCG 658
97 gIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeuIleA 114
659 CCAGCTTGGCTCTGGGAGACCTGCTGCACATCGTCATTGACATCCCTATC 708
114 laSerLeuAlaLeuGlyAspLeuIleTyrIleValIleAspIleProIle 130
709 AATGCTCTACAAGCTGCTGCGACAGGAGCTGCCCATTTGGAGCTGAG..... 753


```
|||||
131 IleValTyrLysLeuLeuAlaGlnLysTrpProPheGlyAspSerGluPh 147
754 .....ATGCTAGCTGGTCCCTTTCATACAGAAAGCCTCGGTGG 793
147 eGlyGlnPheLeuCysLysPheLeuProPheIleGlnLysAlaSerValG 164
794 GAATCAGCTGCTGAGTCTATGCTGCTGAGTATTGACAGATATCGAGCT 843
164 lYleThrValLeuAsnLeuCysAlaLeuSerValAspArgTyrArgAla 180
844 GTTGCTCTTGGAGTAGAATTAAAGGAATGGGGTTCACAAATGGACAGC 893
181 ValAlaSerTrpSerArgValGlnGlyIleGlyIleProMetIleThrAl 197
894 AGTAGAAATGTTTATGATTGGGTGCTCTGCTGCTGCTGCTGCTGCTG 943
197 aileGluIlePheSerIleTrpLeuSerPheIleLeuAlaIleProG 214
944 AAGCCATAGGTTTTCATATATACGATGAGTACAAAGGAAGTATCTG 993
214 luAlaIleGlyPheAlaValValProPheArgTyrLysAspGluSerTyr 230
994 CGAATCTGCTGCTTCATCCCTTCAGAGACAGAGCTTTCATGAGTTTA 1043
231 ValThrCysMetLeuAsnProThrAsnLys.....PheMetLeuPheTy 245
1044 CAAGACAGAAAGATGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1093
245 rLysAspAlaLysAspTrpTrpLeuPheGlyPheTyrPheCysMetProL 262
1094 TGGCCATCACTGCATTTTATACACTAATGACCTGTGAATGTTG... 1140
262 euAlaCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuAsn 278
1141 AGAAAGAAAGTGGCAGATGCTTTAAATGATGACCTTAAGCAGAG 1190
279 ArgArgAsnSerAsnLeuArgIleAlaLeuSerGluHisLeuLysGlnAr 295
1191 ACGGAGTGGCCAAACCGCTTTTGGCTGGTCTGCTGCTGCTGCTGCTG 1240
295 gArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLeu 312
1241 GCTGGCTTCCCTTCACTACCTCAGCAGATTCCTGAAGCTCACTCTTTA 1290
312 ysrTrpPheProLeuHisLeuSerArgIleLeuLysLysMetValTyrAsn 328
1291 CAGAAATGATCCCAATAGATGTGACATTTTGGAGCTTTCTGTTGATGGA 1340
329 GluArgAspProGlyArgCysGluLeuLeuSerPheLeuProLeuAs 345
1341 CTATATTGTATCAACATGGCTTCATGATTCCTGCTGATTAACCCCAATTG 1390
345 pTyrIleSerIleAsnLeuAlaThrMetAsnSerCysIleAsnProIleA 362
1391 CTCTGATTGTTGGTGAAGAAAGATTCAAAAGCTGCTTTAAGTCATGCTTA 1440
362 laLeuTyrPheValSerLysLysPheLysAsnLysPheGlnSerCysLeu 378
1441 TGCTGCTGGTGC...CAGTCATTTGAAGAAACAGTCCCTTGGAGGAAA 1487
379 CysCysCysCysSerGlnSerLysSerLeuAlaThrSerValProMetAs 395
1488 GCAGTCGCTGCTTAAAGTTCAAA.....GCTAATGATCAGGATATG 1528
395 nGlyThrSerIleGlnTrpLysAsnGlnGluLeuAsnAsnHisAsnThrA 412
1529 ACAACTTCGTTCCAGTAAATAATACAGC 1557
412 sp.....ArgSerSerHisLysAspSer 419
seq_name: sp_vertibrate:091548
```

```
seq_documentation_block:
ID Q91548 PRELIMINARY; PRT; 415 AA.
AC Q91548;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE ENDOTHELIN RECEPTOR.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=HEART;
RX MEDLINE=94230448; PubMed=8173772;
RA Kumar C., Wangi V., Nuthulaganti P., Wu H.L., Pullen M., Brun K.,
RA Aliyar H., Morris R.A., Naughton R., Nambi P.;
RT "Cloning and characterization of a novel endothelin receptor from
RT Xenopus heart.";
RL J. Biol. Chem. 269:13414-13420(1994).
DR EMBL; U06633; AAA19570.1;
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 415 AA; 47003 MW; 98DA919097085D70 CRC64;
```

```
alignment_scores:
Quality: 1244.50 Length: 428
Ratio: 3.760 Gaps: 6
Percent Similarity: 77.336 Percent Identity: 56.542
```

```
alignment_block:
```

```
US-09-931-157-2 x Q91548 ..
```

```
Align seg 1/1 to: Q91548 from: 1 to: 415
```

```
265 CGCGCCCTGGTGGTCTGCTTCTTGGC...TGGGCGCTGTGCGGATCTG 311
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
7 ArgPheThrValLeuLeuValLeuAlaGlyIleAlaValSerSerPh 23
312 GGGAGAGGAGAGAGGCTTCCCGCTGACAGGCGCACTCCGCTTTGCAAA 361
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
23 eGlyGluTyrTyrGlnAsnArgThrAspAlaSerThrAspPheThrMetL 40
362 CGCAGAGATAATGACGCCACCCACTAAGACCTTATGCCCCAAGGTTCC 411
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
40 euAsnArgSerHisThrSerPro.....CysProG 47
412 AACGCCAGTCTGGCGCGTCTGTTGGCAGCTGCGGAGGTGCCTAAAGGAGA 461
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
48 .....ValArgLysGlyAs 52
462 CAGGACGGCAGGATCTCCGCCACGACCATCTCCCTCCCGCGTGGCCAA 511
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
52 nArgSerAlaAspLeu.....CysProG 60
512 GACCCATCGAGATCAAGAGACTTTCATATACATCAACACGCTTGTGTCC 561
:||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 lLysThrLysIleAsnHisValPheLysTyrIleAsnThrIleLeuSer 76
562 TGCTTGTGTCGCTGGGATCATCGGAACCTCCACACTTCTCTGAGAA 611
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
77 CysThrIlePheIleGlyMetValGlyAsnAlaThrLeuLeuArgI 93
612 TATCTACAAGAAAGTGCATGCGAAACGGTCCCAATATCTTGATCGCA 661
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
93 eLleTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeuIleAla 110
```

```
662 GCTTGGCTGGGAGACCTCTGCACATCGTCATTCGACATCCCTATCAAT 711
110 erLeuAlaLeuGlyAspLeuIleThrValIleAspIleProIle 126
712 GTCTACAAGCTGCTGGCAGAGGACTGCCATT.....GG 746
127 ValPheLeuLeuGlyGlnArgTrpPheAspGlnSerProValG 143
747 AGCTGAGTGTGAAGCTGCTGCTTCATACAGAAAGCCCTCGTGGA 796
143 YAlaPheLeuCysLysLeuValProPheIleGlnLysAlaSerValGly 160
797 TCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 846
160 leThrValLeuAsnLeuCysAlaLeuSerValAspArgTrpArgAlaVal 176
847 GCTCTTGGAGTGAATTAAGAAGTGGGGTCCAAATGGACAGCAGT 896
177 ProSerTrpSerArgValGlnGlySerGlyIleProLeuIleThrAla 193
897 AGAATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 946
193 eGluIleSerIleTrpValLeuSerPheValLeuAlaIleProGlu 210
947 CCATAGTGTGATATTAATACGATGACGATACAAAGAACTTATCTCG 996
210 laIleGlyPheValMetValProPheGluTrpArgGlyGluGlnPhe 226
997 ATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1046
227 ThrCysMetPheIleAlaThr.....SerProPheMetMetPheTrp 241
1047 GACAGCAAAAGATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1096
241 sAsnAlaLysAspTrpLeuPheGlyLeuTrpPheCysValProLeu 258
1097 CCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
258 laCysThrGlyValPheTrpThrMetMetThrCysGluMetLeuHis 274
1147 AAAAGTGGC...ATGCAGATTGCTTTAAATCATCACCTAAAGCAG 1193
275 ArgGlySerLeuArgIleAlaLeuSerGluHisLeuLysGlnArg 291
1194 GGAAGTGGCCAAACCGCTCTTTGCCGCTGCTGCTGCTGCTGCTGCT 1243
291 gGluValAlaLysThrValPheCysLeuValValIlePheAlaLeu 308
1244 GGCTTCCCTTCCCTCAGCAGGATTCGAGCTCACTCTTATATACAG 1293
308 rpPheProLeuHisLeuSerArgIleIleLysAsnThrIleTrpAsn 324
1294 AATGATCCCAATAGATGTAACCTTTTCAGCTTTCTGTTGTTGTA 1343
325 LeuAspProSerArgCysGluLeuLeuSerPheLeuLeuValMetAsp 341
1344 TATTGGTATCAACATGCTTCACTGAATTCCTGCTGCTGCTGCTGCT 1393
341 eIleSerIleAsnLeuAlaAlaLeuAsnSerCysIleAsnProIleAl 358
1394 TGTATTGGTGAGCAAAAGATTCAAACTGCTTTTAAGTATGCTTATGC 1443
358 eutyPheValSerLysLysPheLysAsnCysPheGlnSerCysLeu 374
1444 TGCTGGTCCCACTATTGTAAGAAACACAGTCTTGGAGGAAAGCAGTC 1493
375 CysCysCysGlnSerLysThrHisIleAsnThrAlaProMetAsnVal 391
1494 GTGCTTAAAGTTCAAGCTAATGATCACGGATAT 1527
391 rSerIleGlnTrpLysAsnHisAspGlnAsnTrp 402
```

seq_name: sp_rodent:091VV2

```
seq_documentation_block:
ID Q91VV2 PRELIMINARY; PRT; 427 AA.
AC Q91VV2;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE SIMILAR TO ENDOTHELIN RECEPTOR TYPE A.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BREAST TUMOR;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC008277; AA08277.1;
KW Receptor.
SQ SEQUENCE 427 AA; 48578 MW; 10142BE73B028233 CRC64;

alignment_scores:
Quality: 1239.00 Length: 380
Ratio: 3.946 Gaps: 7
Percent Similarity: 82.632 Percent Identity: 61.579

alignment_block:
US-09-931-157-2 x Q91VV2 ..
Align seg 1/1 to: Q91VV2 from: 1 to: 427

466 ACAGCAGGATCTCGCCACGACCATCTCCCTCCCGG..... 504
|||||
50 ThrThrHisArgProProAsnLeuAlaLeuProSerAsnGlySerMetHi 66
505 .....TGCCAGGACCCATCGAGATCAAGGAGACTTCAATACATCA 547
|||||
66 sGlyTrpCysProGlnGlnThrLysIleThrThrAlaPheLysTrpIleA 83
548 ACAGGTTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 597
|||||
83 snThrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAla 99
598 ACATCTCTGAGAAATTATCTACAGACAACTGCATGCGAAACGCTCCAA 647
|||||
100 ThrLeuLeuArgIleIleTrpGlnAsnLysCysMetArgAsnGlyProAs 116
648 TATCTTGCATCGCCAGCTTGGCTCTGGGAGACCTGCTGCACATCGTCAT 697
|||||
116 nAlaLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTrpValIleA 133
698 ACATCCCTATCAATGCTACAAAGCTGCTGGCAGAGGACTGGCCA..... 741
|||||
133 spLeuProIleAsnValPheLysLeuLeuAlaGlyArgTrpPropheasp 149
742 .....TTGGAGCTGAGATGCTGAAGCTGCTGCTGCTGCTGCTGCTGCT 782
|||||
150 HisAsnAspPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLy 166
783 AGCTTCGCTGGGAATCAGTCTGCTGAGTCTATGCTGCTGCTGCTGCTGCT 832
|||||
166 sSerSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspA 183
833 GATATCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 882
|||||
183 rgTrpArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIlePro 199
883 AAATGGACAGCAGTGAATAATGTTTGTGTTGTTGTTGTTGTTGTTGTT 932
|||||
200 LeuIleThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLe 216
933 GGCTGTCCCTGAAGCCATAGGTTTGTGATATTAATACGATGACGATACA 982
```


68 lIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThrLeuL 85

605 TGAGAAATTATCTACAAGAACAAAGTCATGCGCAACGGTCCCAATATCTGTG 654

85 euArgIleIleYrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeu 101

655 ATCGCCAGCTTGGCTCTGGAGACCTGCTGCACATCGTCATGGTCATGACATCC 704

102 IleAlaSerLeuAlaLeuGlyAspIleIleYrValIleAspIleuPr 118

705 TATCAATGCTACAAAGCTGCTGGCAGAGACCTGGCCA..... 741

118 orIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheAspHisAsnA 135

742 ..TTTGGAGCTGAGATGTGAAGTGGTGCTCTTCATACAGAAAGACCTCC 789

135 spPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSerSer 151

790 GTGGGAATCACGTGCTGAGTCTATGCTCTCGAGTATTTGACAGATATCG 839

152 ValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTrpAr 168

840 AGCTGTGCTCTTGGAGTAGAATTAAAGGAATTTGGGGTTCCAAAATGA 889

168 gAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeuIleTr 185

890 CAGCAGTAGAAATGTTTTGATTTGGTGGTCTCTCTGCTGTCTGGCTGTC 939

185 hrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAlaIle 201

940 CTTGAGCCATAGCTTTTTCATATAATTACATGGACTACAAAGGAAGTTA 989

202 ProGluAlaIleGlyPheValMetValProPheGluTrpArgGlyGluGl 218

990 TCTGCCAATCTGCTGCTCATCCCTCTCAGAAAGACAGCTTTTCATCGAT 1039

218 nHisLysThrCysMetLeuAsnAlaThrSerLys.....PheMetGluP 233

1040 TTTTACAAGACAGCAAAAGATTGGTGGCTGTTCAGTTTCTATTTCCTGCTG 1089

233 heIyrGlnAspValLysAspTrpLeuPheGlyPheYrPheCysMet 249

1090 CCATTGGCCATCATCGATTTTATACACTTAATCACCTGTGCAATGTT 1139

250 ProLeuValCysThrAlaIlePheYrThrLeuMetThrCysGluMetLe 266

1140 GAGAAAGAAAAGTGGC...ATGCAGATTGCTTTTAAATGATCACCTAAAGC 1186

266 uAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysC 283

1187 AGAGACGGGAAGTGGCCAAACCGTCTTTTGCCTGCTCTGCTGTGCTTGGC 1236

283 lnArgGluValAlaLysThrValPheCysLeuValValIlePheAla 299

1237 CTCTGCTGCTTCCCTTACCTCAGCAGAGATTCTGAAGCTCACTCTTTA 1286

300 LeuCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTy 316

1287 TAATCAGAANTGATCCCAATAGATGCTGAACCTTTTGAGCTTCTGTGGTAT 1336

316 rAspGluMetCAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuL 333

1337 TGAGCTATTGTTATCAACTGCTTCACTGAATTCCTGCATTAACCCA 1386

333 etAspTrpIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnPro 349

1387 ATTGCTCTGATTGTTGGACAAAGATTCAAACCTGCTTTTAAAGTCATG 1436

350 IleAlaLeuYrPheValSerLysLysPheLysAsnCysPheGlnSerCy 366

1437 CTTATGCTGCTGGTGC 1452

1466 sLeuCysCysCysCys 371

1287 TAATCAGAATGATCCCCAATAGACTGTGAACCTTTTGAGCTTCTGTGGTAT 1336
 1336 rAspGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuM 333
 1337 TGGACTATATGGTATCAACATGGCTTCACTGAATTCCTGCATTAAACCCA 1386
 1386 etAspTyrIleGlyIleAsnLeuAlaPheMetAsnSerCysIleAsnPro 349
 1387 ATTGCTCTGTATTGGTGACCAAAAGATTCAAACACTGCTTTTAAGTCATG 1436
 1436 IleAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCys 366
 1437 CTTATGCTGCTGGTGC 1452
 1452 sLeuCysCysCysCys 371

seq_name: sp_human:Q16433

```
seq_documentation_block:
ID   Q16433      PRELIMINARY;          PRT;      318 AA.
AC   Q16433;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE   01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE   ENDOTHELIN-A RECEPTOR DELTA 3-4.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96190719; PubMed=8611157;
RA   Miyamoto Y., Yoshimasa T., Arai H., Takaya K., Ogawa Y., Itoh H.,
RA   Nakao K.;
RT   "Alternative RNA splicing of the human endothelin-A receptor generates
RT   multiple transcripts.";
RL   Blochem. J. 313:795-801(1996).
DR   EMBL; S81545; BAB36327.1; -.
DR   InterPro: IPR000276; GPCR_Rhodpsn.
DR   Pfam; PF00001; 7tm_1; 2.
DR   PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 2.
KW   Receptor.
KW   Receptor.
SQ   SEQUENCE 318 AA; 36488 MW; 0423FAA16DB6A9BF CRC64;
```

alignment_scores:
Quality: 838.50 Length: 354
Ratio: 3.794 Gaps: 4
Percent Similarity: 62.429 Percent Identity: 46.328

alignment_block:

US-09-931-157-2 x Q16433 ..

Align seg 1/1 to: Q16433 from: 1 to: 318

```
505 TGCCAGAGACCATCGACATCAGACGACTTCAAAATACATCAACACGGT 554
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
69 CysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnThrVa 85
555 TGTGTCCTCGCTGTTGCTGCTGGGATCATCGGAACCTCCACACTTC 604
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
85 IleserCysThrIlePheIleValGlyMetValGlyAsnAlaThrLeuL 102
605 TGAGAAATATCTACAGAACAAAGTCATCGCAACGGTCCCAATATCTTG 654
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
102 euArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeu 118
655 ATCGCCAGCTGGCTCTGGGAGACCTGTCGCATCGTCATTCATGACATCCC 704
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
119 IleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspLeuPr 135
705 TATCAATGCTACAAAGCTCTGCGCAGAGGACTGGCCATTGGAGCTGAGA 754
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
135 oIleasnValPheLys..... 140
755 TGTGTAAGCTGGTGCCTTTTCATACAGAAAGCCTCGGTGGAAATCACTGTG 804
140 ..... 140
805 CTGAGTCTATGCTCTGCTGAGTATTGACAGATATCGAGCTGTTGCTCTTG 854
140 ..... 140
855 GAGTAGAATTAAAGAAATGGGTTCCAAAATGGACAGCAGTAGAANAATG 904
140 ..... 140
905 TTTTGATTGGTGCTCTGCTGTGGTTCTGGCTGTCCCTGAAGCCATAGGT 954
```

```
140 ..... 140
955 TTTGATATAATTACGATGGACTACAAGAGTATTATCTGGAATCTGCTT 1004
140 ..... 140
1005 GCTTCATCCGTTTCAGAGACAGACGCTTTCATGCGATTTCACAGACAGCAA 1054
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
141 ..... PheTyrGlnAspValL 146
1055 AAGATTGGTGGCTGTTCAGTTCTATTCTGCTGCTGCTGCTGCTGCTGCT 1104
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
146 ysAspTyrTrpLeuPheGlyPheTyrPheCysMetProLeuValCysThr 162
1105 GCATTTTATATACACTAATGACCTGTAATCTTCTGAGAAAGAAAGTGG 1154
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
163 AlaIlePheTyrThrLeuMetThrCysGluMetLeuAsnArgAsnGln 179
1155 C...ATGCAGATTGCTTTAAATGATCACCTAAAGCAGACGAGGAAAGTGG 1201
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
179 ySerLeuArgIleAlaLeuSerGluHisLeuLysGlnArgGluValA 196
1202 CCAAAACCGCTTTTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCC 1251
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
196 laLysThrValPheCysLeuValValIlePheAlaLeuCysTyrPhePro 212
1252 CTTCACTTCAGCAGGATTCGAGCTCACTCTTTAATCAATCAGATGATCC 1301
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
213 LeuHisLeuSerArgIleLeuLysTyrThrValTyrAsnGluMetAspLy 229
1302 CAATAGATGTGAACCTTTGAGCTTTCTGTTGGTATTGGACTATATGGTA 1351
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
229 sasArgCysGluLeuLeuSerPheLeuLeuLeuMetAspTyrIleGlyI 246
1352 TCAACATGGCTTCACGTAATCTGCTGATTAACCAATTCCTGCTGTAFTTG 1401
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
246 leAsnLeuAlaThrMetAsnSerCysIleAsnProIleAlaLeuTyrPhe 262
1402 GTGACAAAGATTCAAAACCTGCTTTAAGTCATGCTTATGCTGCTGCTG 1451
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
263 valSerLysLysPheLysAsnCysPheGlnSerCysLeuCysCysCys 279
1452 C...CAGTCATTTGAAGAAAACAGCTCCTTGAGAGAAAAGCAGCTGCTG 1498
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
279 styrgLysSerLysSerLeuMetThrSerValProMetAsnGlyThrSerI 296
1499 TAAAGTTCAAAGCTAATGAT...CACGGATATGACAACTCCGTTCCAGT 1545
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
296 leGlnTrpLysAsnHisAspGlnAsnAsnHisAsnThrAspArgSerSer 312
1546 AATAAATACAGC 1557
313 HisLysAspSer 316
```

seq_name: sp_mammal:Q9BDJ8

seq_documentation_block:

```
ID   Q9BDJ8      PRELIMINARY;          PRT;      153 AA.
AC   Q9BDJ8;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT   01-OCT-2001 (TrEMBLrel. 18, Last annotation update)
DE   ENDOTHELIN B RECEPTOR (FRAGMENT).
GN   ETER.
OS   Ovis aries (Sheep).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovidae; Caprinae; Ovis.
OX   NCBI_TaxID=9940;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Clark K.E., McElvly S., Mershon J.L.;
```

RT "Mechanism of uterine vascular refractoriness to endothelin-1 in
RT pregnant sheep."
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF349439; AAK17896.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 153
FT NON_TER 153 153
SQ SEQUENCE 153 AA; 18029 MW; D4DFD3D094462970 CRC64;

alignment_scores:
Quality: 805.00 Length: 153
Ratio: 5.261 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 96.078

alignment_block:

US-09-931-157-2 x Q9BDJ8 ..

Align seg 1/1 to: Q9BDJ8 from: 1 to: 153

1000 TGTCTGCTTCATCCCTTCAGAGACAGCTTTCATGCAGTTTACAGAC 1049
|||||
1 CysLeuLeuHisProthrGlnLysThrAlaPheMetGlnPheTyrLysTh 17
1050 AGCAAAAGATTGGTGGCTGTTCAAGTTCTATTCTGCTGCCATTGGCCA 1099
|||||
17 rAlaLysAspTrpTrpPheSerPheTyrPheCysLeuProLeuAlaI 34
1100 TCACTGCATTTTTATACACTAATGACCTGTGAAATGTGAGAAAGAAA 1149
|||||
34 lethrAlaLeuPheTyrThrLeuMetThrCysGluMetLeuArgLysLys 50
1150 AGTGGCATCGAGATTGCTTTAATGATCACCCTAAAGCAGACGGGAAGT 1199
|||||
51 SerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArgArgGluVa 67
1200 GGCAAAACCGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1249
|||||
67 lAlaLysThrValPheCysLeuValLeuValPheAlaLeuCysTrpLeup 84
1250 CCCTTCACCTCAGCAGGATTCTGAAGCTCACTCTTTATATATCAAGATGAT 1299
|||||
84 roLeuHisLeuSerArgIleLeuLysLeuThrLeuTyrAspGlnHisAsp 100
1300 CCCAATAGATGTGAACCTTTTGACCTTCTGTTGGTATGGACTATATTGG 1349
|||||
101 ProArgArgCysGluPheLeuSerPheLeuValLeuAspTyrIleG1 117
1350 TATCAACATGGCTTCACTGAATTCCTGCATTAACCCAAATGCTGTGATT 1399
|||||
117 yIleAsnMetAlaSerLeuAsnSerCysIleAsnProIleAlaLeuTyrL 134
1400 TGGTGACAAAGATTCAAAACCTGCTTTAAGTCATGCTTATGCTGCTGG 1449
|||||
134 euValSerLysArgPheLysAsnCysPheLysSerCysLeuCysCysTrp 150
1450 TGGCAGTCA 1458
|||||
151 CysGlnSer 153

seq_name: sp_mammal:Q28467

seq_documentation_block:

ID Q28467 PRELIMINARY; PRT; 176 AA.
AC Q28467;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE ETA RECEPTOR (FRAGMENT).

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC Cercopithecinae; Macaca.
RN NCB1_TaxID=9541;
RP SEQUENCE FROM N.A.
RC TISSUE=LUNG;
RA Hino A., Tokuyama Y., Kobayashi M., Yano M., Weir B., Takeda J.,
RA Bell G.I., Macdonald R.L.;
RT "Increased expression of endothelin B receptor mRNA following
RT subarachnoid hemorrhage in monkeys."
RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases.
DR EMBL; U20577; AAA62436.1; -
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON_TER 1 176
FT NON_TER 176 176
SQ SEQUENCE 176 AA; 20756 MW; E46911A2F8F81484 CRC64;

alignment_scores:
Quality: 667.50 Length: 171
Ratio: 4.279 Gaps: 2
Percent Similarity: 91.228 Percent Identity: 68.421

alignment_block:

US-09-931-157-2 x Q28467 ..

Align seg 1/1 to: Q28467 from: 1 to: 176

943 GAAGCATAGTTTTCATATATATAGATGAGTACAAAGAGATTCATCT 992
|||||
1 GluAlaIleGlyPheValMetValProPheGluTyrArgGlyGluGlnHi 17
993 GCAGATCTGCTTCTTCATCCCTTCAGAGACAGCTTTCATGCAGTTT 1042
|||||
17 sLysThrCysMetLeuAsnAlaThrSerLys.....PheMetGluPheT 32
1043 ACAAGACAGCAAGATTGGTGGCTGTTCAGTTTCTTATTTCTGCTGCCA 1092
|||||
32 yrGlnAspValLysAspTrpTrpLeuPheGlyPheTyrPheCysMetPro 48
1093 TTGCCCATCACTGCATTTTATACACTAATGACCTGTGAAATGTTGAG 1142
|||||
49 LeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuAs 65
1143 AAAGAAAAGTGGC...ATGCAGATTGCTTTAAATGATCACCTAAAGCAGA 1189
|||||
65 nArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysGlnA 82
1190 GACGGGAAGTGGCAAAACCGCTCTTTTGGCTGGTCTTGTCTTTCGCCCTC 1239
|||||
82 rgArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLeu 98
1240 TGTGTGCTTCCCTTCACCTCAGCAGGATTCTGAAGCTCACTCTTTATAA 1289
|||||
99 CysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyrAs 115
1290 TCAGAAATGATCCCAATAGATGTGAACCTTTTGAGCTTTCTGTTGCTATTGG 1339
|||||
115 nGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeuMetA 132
1340 ACTATATTGGTATCAACATGCTTTCACCTGAATTCCTGATTAACCCCAATT 1389
|||||
132 spTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnProIle 148
1390 GCTCTGTATTTGGTGACAAAGATTCAAAACCTGCTTTAAGTCATGCTTT 1439
|||||
149 AlaLeuTyrPheValSerLysLysLysPheLysAsnCysPheGlnSerCysLe 165

556 GTGTCCTGCCCTGTGCTCGTGGGATCATCGGAACTCCACACTTCT 605
100 ValSerCysLeuValPheValLeuGlyIleIleGlyAsnSerThrLeuLe 116
606 GAGAAATTATCTACAAGAACAGTGCATCGGAACCGGTCCCAATATCTTGA 655
116 uArgIleIleIleIleIleIleIleIleIleIleIleIleIleIleIleIle 133
656 TCGCCAGCTTGGCTCTGGGAGAC 678
133 leAlaSerLeuAlaLeuGlyIleIleIleIleIleIleIleIleIleIleIle 140

```

1440 ATGCTGCTGCTGC 1452
|||||
165 uCysCysCysCys 169

seq_name: sp_rodent:Q9JL53
seq_documentation_block:
ID Q9JL53 PRELIMINARY; PRT; 140 AA.
AC Q9JL53;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE ENDOTHELIN-B RECEPTOR (FRAGMENT).
DE EDRB.
GN Peromyscus maniculatus (Deer mouse).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Sigmodontinae;
OC Peromyscus.
OX NCBI_TaxID=10042;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=BAIRDII;
RA Vrana P.B., Fossella J.A., Matteson P., del Rio A., O'Neill M.J.,
RA Tilghman S.M.;
RT "Genetic and Epigenetic Incompatibilities Underlie Hybrid Dysgenesis
RT in Peromyscus.";
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF212999; AAF62405.1; -.
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm1.1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
DR Receptor.
KW 1
FT NON_TER 140
FT NON_TER 140
SQ SEQUENCE 140 AA; 15050 MW; F0107E422AEF4AE2 CRC64;

```

```

alignment_scores:
  Quality: 583.00      Length: 141
  Ratio: 4.555        Gaps: 2
  Percent Similarity: 90.780  Percent Identity: 80.851

alignment_block:
  US-09-931-157-2 x Q9JL53

  Align seg 1/1 to: Q9JL53 from: 1 to: 140

259  TCGCGACGCGCCTGGTTCGGCTGGTTCCTGCGGCGCTGTCGCGGAT 308
      |||||
1   CysGlyArgAlaLeuValAlaLeuLeuAlaCysGlyLeuLeuGluVal 17
      |||||

309  CTGGGGAGGAGGAGAGGCTTCGCGCTGCAGGGCCACTCG...CTTT 355
      |||||
17  ltrpGlyGluGluArgGlypheProProAlaGlnAlaThrProSerLeuL 34
      |||||

356  TGCAAACCCGACGAGATAATGACGCCACCCACTAAGACCTTATGGCCCAAG 405
      |||||
34  euGlyThrArgGluValMetThrProProThrLysThrSerTrpThrArg 50
      |||||

406  GGTTCACAGCCAGCTGCGCGCGGTGCTTGGCACCTCGGAGGTGCCPAA 455
      |||||
51  GlySerAsnSerSerLeuThrArgSerLeuAlaProAlaGluValProLy 67
      |||||

456  AGGACACAGGACGGCAGGATCTCCGCGACACCATCTCCCTCCGCCGT 505
      |||||
67  sGlyGlyArgValAlaGlyValArgProArgSerPhe...ProProProC 83
      |||||

506  GCCAAGGCCCATCGAGATCAAGGAGACTTTCAATACATCAACACGCTT 555
      |||||
83  ysGlnArgAsnIleGluIleSerLysThrPheLysTyrIleAsnThrIle 99
      |||||

```

OM of: US-09-931-157-2 to: SwissProt_40:* out_format : pfs

Date: Oct 1, 2002 1:21 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame*_n2p.model -DEV=xlp
-O=/cgn2.1/USPTO.spool/US09931157/runat_30092002.110421.11329/app_query.fasta_1.8657
-DB=SwissProt_40 -OPMT=fastan -SUFFIX=rs -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -START=1 -MATRIX=blosum62
-DELOP=6.000 -DELEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-TRANS=human40.cdi -LIST=5 -DOALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09931157 -CGN1_1_127 -NCPU=6 -ICPU=3 -LONGLOG
-DEV=TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-931-157-2

Query length: 4301

Database: SwissProt_40:*

Database sequences: 105224

Database length: 38719550

Search time (sec): 126.310000

score_list:

Sequence	Strd Orig	ZScore	Escore	Len	Documentation
SwissProt_40:ETBR_HUMAN	+ 2240.00	3295.07	1.7e-176	442	! P24530 homo sapiens (human).
SwissProt_40:ETBR_CANFA	+ 2206.00	3105.87	6.2e-166	426	! P56497 canis familiaris (dog).
SwissProt_40:ETBR_PIG	+ 2121.50	2985.97	2.8e-159	443	! P35463 sus scrofa (pig). endot
SwissProt_40:ETBR_HORSE	+ 2085.50	2935.04	2.0e-156	443	! P62709 equus caballus (horse).
SwissProt_40:ETBR_RAT	+ 2070.00	2913.14	3.3e-155	442	! P24451 rattus norvegicus (rat).
SwissProt_40:ETBR_MOUSE	+ 2068.00	2910.31	4.7e-155	442	! P48302 mus musculus (mouse).
SwissProt_40:ETBR_BOVIN	+ 2046.50	2879.92	2.3e-153	441	! P28088 bos taurus (bovine).
SwissProt_40:ETBR_COTJA	+ 1705.50	2399.83	1.6e-126	347	! P90328 coturnix coturnix jap
SwissProt_40:ETBR_RAT	+ 1246.50	1748.60	2.5e-90	426	! P26684 rattus norvegicus (rat).
SwissProt_40:ETBR_HUMAN	+ 1240.00	1739.38	8.1e-90	427	! P25101 homo sapiens (human).
SwissProt_40:ETBR_BOVIN	+ 1238.00	1736.55	1.2e-89	427	! P21450 bos taurus (bovine).
SwissProt_40:ETBR_PIG	+ 1233.00	1729.48	2.9e-89	427	! Q29010 sus scrofa (pig). endot
SwissProt_40:ETBR_XENLA	+ 1073.00	1502.78	1.2e-76	444	! P32940 xenopus laevis (afri
SwissProt_40:ETBR_MOUSE	+ 1019.00	1431.12	1.9e-72	270	! P61614 mus musculus (mouse).
SwissProt_40:ETBR_MACFA	+ 504.00	712.16	5.7e-32	99	! Q28468 macaca fascicularis (c
SwissProt_40:GRPR_MOUSE	+ 498.50	691.50	2.1e-31	384	! P21729 mus musculus (mouse).
SwissProt_40:GRPR_RAT	+ 496.50	688.67	3.0e-31	384	! P52500 rattus norvegicus (rat).
SwissProt_40:NNBR_HUMAN	+ 491.00	680.74	8.2e-31	390	! P28336 homo sapiens (human).
SwissProt_40:NNBR_MOUSE	+ 490.50	680.03	8.9e-31	390	! P54799 mus musculus (mouse).
SwissProt_40:NNBR_RAT	+ 489.50	678.62	1.1e-30	390	! P24053 rattus norvegicus (rat).
SwissProt_40:GRPR_HUMAN	+ 466.00	645.52	7.6e-29	384	! P30550 homo sapiens (human).
SwissProt_40:BRSS_MOUSE	+ 448.00	619.70	2.0e-27	399	! P54798 mus musculus (mouse).
SwissProt_40:BRSS_SHEEP	+ 447.50	618.99	2.2e-27	399	! P97967 ovis aries (sheep). B
SwissProt_40:BRSS_BOMOR	+ 443.50	613.50	4.5e-27	399	! P47751 bombina orientalis (c
SwissProt_40:BRSS_HUMAN	+ 443.00	612.63	5.0e-27	399	! P32247 homo sapiens (human).
SwissProt_40:BRSS_CAVPO	+ 441.00	609.80	7.1e-27	399	! P35371 cavia porcellus (guin
SwissProt_40:GP37_HUMAN	+ 406.00	556.21	4.5e-24	613	! P15354 homo sapiens (human).
SwissProt_40:EBP2_HUMAN	+ 387.00	531.63	1.3e-22	481	! P60883 homo sapiens (human).
SwissProt_40:TLRI_DRONE	+ 333.50	455.51	2.2e-18	504	! P30974 drosophila melanogast
SwissProt_40:SSRA_HUMAN	+ 331.50	455.17	3.0e-18	398	! P31391 homo sapiens (human).
SwissProt_40:OPRK_HUMAN	+ 329.50	452.54	4.3e-18	380	! P41145 homo sapiens (human).
SwissProt_40:NY2R_MOUSE	+ 325.00	446.05	9.8e-18	395	! P97295 mus musculus (mouse).
SwissProt_40:SRRL_HUMAN	+ 324.00	444.49	1.2e-17	391	! P30872 homo sapiens (human).
SwissProt_40:NY2R_HUMAN	+ 323.50	444.02	1.3e-17	391	! P49146 homo sapiens (human).
SwissProt_40:SSRL_MOUSE	+ 322.00	441.66	1.7e-17	391	! P30873 mus musculus (mouse).
SwissProt_40:OPRK_CAVPO	+ 321.50	441.22	1.8e-17	380	! P41144 cavia porcellus (guin
SwissProt_40:OPRK_MOUSE	+ 321.50	441.22	1.8e-17	380	! P33534 mus musculus (mouse).
SwissProt_40:NY2R_MACMU	+ 321.50	441.20	1.8e-17	391	! P99k74 macaca mulatta (rhesu
SwissProt_40:OPRK_PIG	+ 318.50	437.88	2.7e-17	401	! Q95247 sus scrofa (pig). mu-ty
SwissProt_40:OPRK_RAT	+ 318.50	436.98	3.2e-17	380	! P34975 rattus norvegicus (rat)

SwissProt_40:NY2R_BOVIN + 318.50 436.88 3.2e-17 384 ! P79113 bos taurus (bo
SwissProt_40:GPRR_HUMAN + 318.00 436.52 3.5e-17 370 ! P49683 homo sapiens (f
SwissProt_40:NY2R_PIG + 316.50 434.10 4.6e-17 382 ! P02836 sus scrofa (pig).
SwissProt_40:SSRA_RAT + 316.50 434.05 4.6e-17 384 ! P30937 rattus norvegicus
SwissProt_40:OPRK_BOVIN + 313.00 428.69 8.7e-17 401 ! P79350 bos taurus (bo
seq_name: SwissProt_40:ETBR_HUMAN
seq_documentation_block:
ID ETBR_HUMAN STANDARD; PRT; 442 AA.
AC P24530; O15433; O9UOK3;
DT 01-MAR-1992 (Rel. 21, Created)
DT 01-MAR-1992 (Rel. 21, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
DE Non-selective type).
GN EDNRB OR ETBR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91254298; PubMed=1710450;
RA Nakamura M., Takayanagi R., Sakai Y., Sakamoto S., Hagiwara H.,
RA Mizuno T., Saito Y., Hirose S., Yamamoto M., Nawata H.;
RT "Cloning and sequence analysis of a cDNA encoding human non-selective
RT type of endothelin receptor.";
RL Biochem. Biophys. Res. Commun. 177:34-39(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91298956; PubMed=1648908;
RA Ogawa Y., Nakao K., Arai H., Nakagawa O., Hosoda K., Suga S.,
RA Nakanishi S., Imura H.;
RT "Molecular cloning of a non-isopeptide-selective human endothelin
RT receptor.";
RL Biochem. Biophys. Res. Commun. 178:248-255(1991).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=93180293; PubMed=1291713;
RA Arai H., Nakao K., Hosoda K., Ogawa Y., Nakagawa O., Komatsu Y.,
RA Imura H.;
RT "Molecular cloning of human endothelin receptors and their expression
RT in vascular endothelial cells and smooth muscle cells.";
RL Jpn. Circ. J. 56:1303-1307(1992).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=93155196; PubMed=8429023;
RA Arai H., Nakao K., Takaya K., Hosoda K., Ogawa Y., Nakanishi S.,
RA Imura H.;
RT "The human endothelin-B receptor gene. Structural organization and
RT chromosomal assignment.";
RL J. Biol. Chem. 268:3463-3470(1993).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE=91315496; PubMed=1713452;
RA Sakamoto A., Yanagisawa M., Sakurai T., Takawa Y., Yanagisawa H.,
RA Masaki T.;
RT "Cloning and functional expression of human cDNA for the ETB
RT endothelin receptor.";
RL Biochem. Biophys. Res. Commun. 178:656-663(1991).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE=93179382; PubMed=8440682;
RA Elshourbagy N.A., Korman D.R., Wu H.L., Sylvester D.R., Lee J.A.,
RA Nurchalaganti P., Bergsma D.J., Kumar C.S., Nambi P.;
RT "Molecular characterization and regulation of the human endothelin
RT receptors.";
RL J. Biol. Chem. 268:3873-3879(1993).
RN [7]
RP SEQUENCE FROM N.A.
RX TISSUE=Lung;
MEDLINE=93132955; PubMed=1282938;

RA Haendler B., Hechler U., Schleuning W.D.;
 RL "Molecular cloning of human endothelin (ET) receptors ETA and ETB.";
 RN J. Cardiovasc. Pharmacol. 20:S1-S4(1992).
 RP [8]
 RC SEQUENCE FROM N.A. (ISOFORM B).
 RN TISSUE=Placenta;
 RA Elahourbagy N., Adamou J., Gagnon A., Wu H., Pullen M., Nambi P.;
 RL "Molecular characterization of a human endothelin receptor splice
 RT variant.";
 RN Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
 RP [9]
 RC SEQUENCE FROM N.A. (ISOFORM DELTA 3).
 RN MEDLINE-99173868; PubMed-10072757;
 RA Tsuchumi M., Liang G., Jones P.A.;
 RL "Novel endothelin B receptor transcripts with the potential of
 RT generating a new receptor.";
 RN Gene 228:43-49(1999).
 RP [10]
 RC PALMITOYLATION.
 RN MEDLINE-97407956; PubMed-92611180;
 RA Okamoto Y., Ninomiya H., Tanioka M., Sakamoto A., Miwa S., Masaki T.;
 RL "Palmitoylation of human endothelinB. Its critical role in G protein
 RT coupling and a differential requirement for the cytoplasmic tail by G
 protein subtypes.";
 RN J. Biol. Chem. 272:21589-21596(1997).
 RP [11]
 RC REVIEW ON VARIANTS.
 RN MEDLINE-98023959; PubMed-9359036;
 RA Hofstra R.M.W., Osianga J., Buys C.H.C.M.;
 RL "Mutations in Hirschsprung disease: When does a mutation contribute to
 RT the phenotype.";
 RN Eur. J. Hum. Genet. 5:180-185(1997).
 RP [12]
 RC VARIANT HSCR CYS-276.
 RN MEDLINE-93094300; PubMed-8001158;
 RA Puffenberger E.G., Hosoda K., Washington S.S., Nakao K., Dewit D.,
 RL Yanagisawa M., Chakravarti A.;
 RT "A missense mutation of the endothelin-B receptor gene in multigenic
 RL Hirschsprung's disease.";
 RN Cell 79:1257-1266(1994).
 RP [13]
 RC VARIANT WS/HSCR GLY-183.
 RN MEDLINE-96177681; PubMed-8634719;
 RA Attie T., Till M., Pelet A., Amiel J., Edery P., Boutrand L.,
 RL Munnich A., Lyonnet S.;
 RT "Mutation of the endothelin-receptor B gene in Waardenburg-
 RT Hirschsprung disease.";
 RN Hum. Mol. Genet. 4:2407-2409(1995).
 RP [14]
 RC VARIANT HSCR ASN-305.
 RN MEDLINE-97005360; PubMed-8852659;
 RA Auricchio A., Casari G., Staiano A., Ballabio A.;
 RL "Endothelin-B receptor mutations in patients with isolated
 RT Hirschsprung disease from a non-inbred population.";
 RN Hum. Mol. Genet. 5:351-354(1996).
 RP [15]
 RC VARIANTS HSCR SER-57; TRP-319 AND LEU-383.
 RN MEDLINE-97005361; PubMed-8852660;
 RA Amiel J., Attie T., Jan D., Pelet A., Edery P., Bidaud C., Lacombe D.,
 RL Tam P., Simeoni J., Flori E., Nihoul-Fekete C., Munnich A.,
 RT Lyonnet S.;
 RT "Heterozygous endothelin receptor B (EDNRB) mutations in isolated
 RL Hirschsprung disease.";
 RN Hum. Mol. Genet. 5:355-357(1996).
 RP [16]
 RC VARIANT HSCR ILE-374, AND VARIANT SER-57.
 RN MEDLINE-96224410; PubMed-8630503;
 RA Hofstra R.M.W., Tan-Sindhunata G., Wu Y., Kamsteeg E.-J., Stulp R.P.,
 RA van Ravenswaaij-Arts C., Majoor-Krakauer D., Angrist M.,
 RA Chakravarti A., Meijers C., Buys C.H.C.M.;
 RT "A homozygous mutation in the endothelin-3 gene associated with a
 RT combined Waardenburg type 2 and Hirschsprung phenotype (Shah-
 RT Waardenburg syndrome).";

RL Nat. Genet. 12:445-447(1996).
 RN [17]
 RP VARIANT HSCR SER-57.
 RX MEDLINE-98430996; PubMed-9760196;
 RA Svensson P.J., Anvret M., Molander M.L., Nordenskjold A.;
 RT "Phenotypic variation in a family with mutations in two Hirschsprung-
 RL related genes (RET and endothelin receptor B).";
 RN Hum. Genet. 103:145-148(1998).
 CC -!- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
 CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND DELTA;
 CC ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- DISEASE: DEFECTS IN EDNRB ARE A CAUSE OF TYPE IV (WS4 OR SHAH-
 CC WAARDENBURG SYNDROME) (WS/HSCR) WHICH IS CHARACTERIZED BY THE
 CC ASSOCIATION OF WS AND HIRSCHSPRUNG DISEASE (HSCR).
 CC -!- DISEASE: DEFECTS IN EDNRB ARE THE CAUSE OF TYPE 2 HIRSCHSPRUNG
 CC DISEASE (HSCR2) (OR AGANGLIONIC MEGACOLON), A CONGENITAL DISORDER
 CC CHARACTERIZED BY ABSENCE OF ENTERIC GANGLIA ALONG A VARIABLE
 CC LENGTH OF THE INTESTINE. HSCR IS THE MOST COMMON RANGE OF
 CC CONGENITAL INTESTINAL OBSTRUCTION EARLY SYMPTOMS FROM
 CC COMPLETE ACUTE NEONATAL OBSTRUCTION, CHARACTERIZED BY VOMITING,
 CC ABDOMINAL DISTENTION AND FAILURE TO PASS STOOL, TO CHRONIC
 CC CONSTIPATION IN THE OLDER CHILD.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; W74921; AAA58465.1; -
 CC EMBL; D90402; BAA14398.1; -
 CC EMBL; L06623; AAB52342.1; -
 CC EMBL; S44866; AAB19411.1; -
 CC EMBL; S57283; AAB25531.1; -
 CC EMBL; D13168; BAA02445.1; -
 CC EMBL; D13162; BAA02445.1; JOINED.
 CC EMBL; D13163; BAA02445.1; JOINED.
 CC EMBL; D13164; BAA02445.1; JOINED.
 CC EMBL; D13166; BAA02445.1; JOINED.
 CC EMBL; D13167; BAA02445.1; JOINED.
 CC EMBL; X99250; CAA67623.1; -
 CC EMBL; AF114165; AAD24541.1; -
 CC PIR; JQ1042; JQ1042.
 CC PIR; B45481; B45481.
 CC PIR; A46609; A46609.
 CC GCRDB; GCR_0068; -
 CC GCRDB; GCR_0402; -
 CC GCRDB; GCR_0474; -
 CC GCRDB; GCR_0615; -
 CC GCRDB; GCR_0701; -
 CC GCRDB; GCR_0728; -
 CC GCRDB; GCR_0736; -
 CC GCRDB; GCR_2423; -
 CC MIM; 131244; -
 CC MIM; 277580; -
 CC MIM; 600155; -
 CC MIM; 142623; -
 CC alignment_scores:
 CC Quality: 2340.00 Length: 442
 CC Ratio: 5.294 Gaps: 0
 CC Percent Similarity: 100.000 Percent Identity: 100.000
 CC alignment_block:
 CC US-09-931-157-2 x ETBR_HUMAN
 CC Align seg 1/1 to: ETBR_HUMAN from: 1 to: 442

```
238 ATGACGCCCTCCCAAGTCTGTGGGAGCGCCCTGGTTGGCGGTCT 287
|||||
1 MetGlnProProSerLeuCysGlyArgAlaLeuValAlaLeuVal 17
288 TGCCTGCCGCTGTCCGCGATCTGGGAGAGGAGAGAGCTCCCGCTG 337
|||||
17 uAlaCysGlyLeuSerArgIleTrpGlyGluGluArgGlyPheProP 34
338 ACAGGGCCACTCCGCTTTTGGAAACCCGAGAGATAATGACGCCACCCACT 387
|||||
34 sPaArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProProThr 50
388 AAGACCTTATGCCCCAAGGGTTCCAAACCCAGCTCTGGCGGCTGTGGC 437
|||||
51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeuAl 67
438 ACCTCGGAGGTGCTTAAGGAGAGACGAGCGCAGGATCTCCGCCACGCA 487
|||||
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArg 84
488 CCATCTCCCTCCCGCTGCCAAGCACCATCGAGATCAAGGAGACTTTC 537
|||||
84 hrIleSerProProCysGlnGlyProIleGluIleLysGluThrPhe 100
538 AAATACATCAACACGGTTGTGTCCTGCTTGTGCTGCTGGGATCAT 587
|||||
101 LysTyrIleAsnThrValValSerCysLeuValPheValLeuGlyIleI 117
588 CGGAACTCCACACTCTGAGAAATATCTACAAGAACAGTGCATCGCGAA 637
|||||
117 eGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArg 134
638 ACGGTCCTCAATCTTGTATCGCCAGCTTGGCTCTGGGAGACTGTGCAC 687
|||||
134 snglyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHis 150
688 ATCGTCATTGACATCCCTATCATCTGTACAAAGCTGCTGGCAGAGGACTG 737
|||||
151 IleValIleAspIleProIleAsnValTyrLysLeuLeuAlaGluAspTr 167
738 GCATTTGGAGCTGAGATGTGTAAGCTGTGCTGCTTCATACAGAAGCT 787
|||||
167 pProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CGTGGGAATCACTGTGCTGACTGTATGTGCTGTGAGTATTCACAGATAT 837
|||||
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
838 CGAGCTGTGCTTCTTGGAGTAGAATTAAGGAATGGGGTTCCAAAATG 887
|||||
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
888 GACACAGTAGAATTTGTTGATTTGGTGGTGTCTGCTGCTGCTGCTG 937
|||||
217 pThrAlaValGluIleValLeuIleTrpValValSerValValLeuAla 234
938 TCCCTGAACCCATAGCTTTTGATATATTAATACCATGACTACAAAGGA 987
|||||
234 aProGluAlaIleGlyPheAspIleIleThrMetAspTyrLysGlySer 250
988 TATCTGCGAATCTGCTTGTCTTCCCGTTCCAGAGACAGCTTTCATGCA 1037
|||||
251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetG 267
1038 GTTTTACAGACAGCAAAAGATTGGTGGCTGTGCTAGTTCTTCTTCTGCT 1087
|||||
267 nPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheCysL 284
1088 TGCCATTGCCATCACTGCTGCTTTTATATACACTATGACTCTGGAATG 1137
|||||
284 euProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGluMet 300
```

```
1138 TTGACAAAGAAAGTGGCATCGCAGATTGCTTTAAATGATCACCCTAAAGCA 1187
|||||
301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
1188 GAGCGGGAAGTGGCCAAACCGCTTTTGGCTGTGCTGTGCTTTGCTTGGCC 1237
|||||
317 nArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 334
1238 TCTGCTGCTTCCCTTCCCTCAGCTCAGCAGATTCTGAAGCTCAGCTTTTAT 1287
|||||
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyr 350
1288 AATCAGATGATCCCAATAGATGTCAACTTTTGACCTTTCTGTTGGTATT 1337
|||||
351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuVal 367
1338 GGACTATATTGGTATCAACATGGCTTCACTGAATTCCTGCATTAACCCAA 1387
|||||
367 uAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnProI 384
1388 TTGCTCTCTATTGCTGAGCAAPAGATTCAAAAACCTGCTTTTAAGTCATGC 1437
|||||
384 IleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCys 400
1438 TTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1487
401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGlu 417
1488 CGAGTCGTGCTTAAAGTTCAAGCTTAATGATCAGGATATGACAACTTCC 1537
|||||
417 sGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnPhe 434
1538 GTTCAGTAATAATACAGCTCATCT 1563
|||||
434 rgSerSerAsnLysTyrSerSerSer 442
```

seq_name: SwissProt_40; ETBR_CANFA

seq_documentation_block:

ID ETBR_CANFA STANDARD; PRT; 425 AA.

AC P56497;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1999 (Rel. 38, Last sequence update)

DT 15-JUL-1999 (Rel. 38, Last annotation update)

DE Endothelin B receptor precursor (ET-B) (Endothelin receptor

DE Non-selective type) (Fragment).

GN EDNRB.

OS Canis familiaris (Dog).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.

OX NCBI_TaxID=9615;

RN (1)

RP SEQUENCE FROM N.A.

RA Zenke D., Yuzbasiyan-Gurkan V.;

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.

CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATES

CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.

CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@sib-sib.ch).

CC EMBL: AF034530; AAC26970.1; -.

CC GCRDB; GCR_2495; -.

CC InterPro: IPR000276; GPCR_Rhodopsn. -.

CC Pfam: PF00001; 7tm_1; 1.

DR PRINTS; PRO0237; GPCR RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECP_F1_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Lipoprotein; Palmitate.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 >426 ENDOPEPTIDASE RECEPTOR.
FT DOMAIN 27 101 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 102 126 1 (POTENTIAL).
FT DOMAIN 127 137 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 138 163 2 (POTENTIAL).
FT DOMAIN 164 175 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 176 197 3 (POTENTIAL).
FT DOMAIN 198 218 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 219 243 4 (POTENTIAL).
FT DOMAIN 244 271 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 272 296 5 (POTENTIAL).
FT DOMAIN 297 324 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 325 350 6 (POTENTIAL).
FT DOMAIN 351 362 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 363 389 7 (POTENTIAL).
FT DOMAIN 390 >426 CYTOPLASMIC (POTENTIAL).
FT DISULFID 174 255 BY SIMILARITY.
FT CARBOHYD 59 59 N-LINKED (GLCNAC...) (POTENTIAL).
FT LIPID 402 402 PALMITATE (POTENTIAL).
FT LIPID 403 403 PALMITATE (POTENTIAL).
FT LIPID 405 405 PALMITATE (POTENTIAL).
FT NON_TER 426 426
SQ SEQUENCE 426 AA; 47876 MW; C5AF1675C30FIP26 CRC64;

alignment_scores:

Quality: 2206.00 Length: 426
Ratio: 5.203 Gaps: 0
Percent Similarity: 99.531 Percent Identity: 97.183

alignment_block:

US-09-931-157-2 x ETBR_CANFA ..

Align seg 1/1 to: ETBR_CANFA from: 1 to: 426

238 ATGCAGCGGCTCAAGTGTGGGAGCGGCCCTGGTGGCGGTCTCT 287
|||||
1 MetGlnProProSerLeuCysGlyArgAlaLeuValAlaLeuVal 17
288 TGCTGGCGGCTGCGCGGATCTGGGAGGAGGAGGCTTCCCGCTG 337
|||||
17 uAlaCysGlyLeuSerArgIleTrpGlyGluArgGlyPheProPro 34
338 ACAGGGCCACTCCGCTTTTGCACACCGCAGAGATAATGACGCCCACT 387
|||||
34 spArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProThr 50
388 AAGACCTTATGGCCAGGTTTCCACGCCAGTCTGGCGGTGGTGGC 437
|||||
51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeu 67
438 ACCTGGCGAGGTGCTAAAGAGACAGGAGGAGGATCTCCGCCACGCA 487
|||||
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArg 84
488 CCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTTC 537
|||||
84 hrIleSerProProCysGluGlySerIleGluIleLysGluThrPhe 100
538 AATACATCAACAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
|||||
101 LysTyrIleAsnThrValValSerCysLeuValPheValLeuGlyIle 117
588 CGGGAATCCACACTTCTGAGAAATATCTACAGAACAAAGTCATGCGAA 637
|||||
117 eGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArg 134

seq_name: SwissProt_40:ETBR_PIG

638 ACGGTCCCAATATCTTGATGCGCCAGCTTGGCTCTGGGAGACCTGCTGCAC 687
|||||
134 snGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHis 150
688 ATCGTCATTGACATCCCTATCAATCTCTACAGCTGCTGGCAGAGACTG 737
|||||
151 IleIleIleAspIleProIleThrValTyrLysLeuLeuAlaGluAspTr 167
738 GCCATTTGGAGCTGAGATGTGTAAGCTGGTGCCTTTTCATACAGAAAGCT 787
|||||
167 pProPheGlyValGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGTGGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 837
|||||
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
838 CGAGCTGTGCTCTTGGAGTAGAATTAAAGGAATGGGTTCCAAAATG 887
|||||
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
888 GACAGCAGTAGAATAATGTTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 937
|||||
217 pThrAlaValGluIleValLeuIleTrpValSerValValLeuAlaVal 234
938 TCCTGGAAGCCATAGTTTGTATATATACGATGAGTACAAAGGAAGT 987
|||||
234 alProGluAlaValGlyPheAspMetIleThrIleAspTrpLysGlyArg 250
988 TATCTGCGAATCTGCTTCTCATCCCGTTCCAGAACAGACAGCTTTCATGCA 1037
|||||
251 TyrLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMetG 1087
1038 GTTTTACAAAGACAGCAAAAGATTGGTGGTGGTGGTGGTGGTGGTGGT 1087
267 nPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheCysL 284
1088 TGCCATTGGCCATCACTGCATTTTATACACTAATGACCTGTGAAATG 1137
|||||
284 euProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGluMet 300
1138 TTGAGAAAGAAAGTGGCATGCAGATTGCTTTAAATGATCACCCTAAAGCA 1187
|||||
301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
1188 GAGCGGGAAGTGGCAAAACCGTCTTTTGGCTGGTGGTGGTGGTGGTGG 1237
|||||
317 nArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaL 334
1238 TCTGCTGGCTTCCCTTCCACCTCAGCAGGATTCGAAAGCTCACTTTTAT 1287
|||||
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrIleTyr 350
1288 AATCAGAAATGATCCCATAGATGCAACTTTTGGCTTCTGCTTCTGCTAT 1337
|||||
351 AspGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuValLe 367
1338 GGACTATATTGCTATCAACATGGCTTCACTGAATTCCTGCATTAACCCAA 1387
|||||
367 uAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnProI 384
1388 TTGCTCTGTATTTGGTGAGCAAAAGATTCAAAACTGCTTTTAAGTCATGC 1437
|||||
384 LeAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCys 400
1438 TTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1487
|||||
401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGlu 417
1488 GCAGTCTGCTTAAAGTTCAAGCTTAAT 1515
417 sGlnSerCysLeuLysPheLysAlaAsn 426

```

seq_documentation_block:
ID  ETBR_PIG STANDARD: PRT: 443 AA.
AC P35463;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
DE Non-selective type).
DE EDNRB.
GN Sus scrofa (Pig).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92186811; PubMed=1312214;
RA Elshourbady N.A., Lee J.A., Korman D.R., Nuthalaganti P.,
RA Sylvester D.R., Dilella A.G., Sutiphong J.A., Kumar C.S.;
RT "Molecular cloning and characterization of the major endothelin
RT receptor subtype in porcine cerebellum.";
RL Mol. Pharmacol. 41:465-473(1992).
CC -!- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
DR GCRdb; GCR 0937; -.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODOPSN.
DR PROSITE; PS00237; G-PROTEIN_RECEP_F1_1; 1.
DR PROSITE; PS00262; G-PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Lipoprotein; Palmitate.
FT SIGNAL 1 26
FT CHAIN 27 443
FT DOMAIN 27 102
FT TRANSMEM 103 127
FT DOMAIN 128 133
FT TRANSMEM 139 164
FT DOMAIN 165 176
FT TRANSMEM 177 198
FT DOMAIN 199 219
FT TRANSMEM 220 244
FT DOMAIN 245 272
FT TRANSMEM 273 297
FT TRANSMEM 298 325
FT TRANSMEM 326 351
FT TRANSMEM 352 363
FT TRANSMEM 364 390
FT DOMAIN 392 443
FT DISULFID 175 256
FT CARBOHYD 60 60
FT CARBOHYD 354 354
FT LIPID 403 403
FT LIPID 404 404
FT LIPID 406 406
FT SEQUENCE 443 AA; 49594 MW; C3B066743990A3C7 CRC64;

```

alignment_scores:

```

Quality: 2121.50 Length: 443
Ratio: 4.992 Gaps: 1
Percent Similarity: 95.937 Percent Identity: 90.068

```

alignment_block:

```

US-09-931-157-2 x ETBR_PIG ..

```

```

Align seg 1/1 to: ETBR_PIG from: 1 to: 443

```

```

238 ATGCAGCCGCTCCAGTCTGTGGAGCGCCCTGGTTCGGTGTCT 287

```

```

|||||
1 MetGlnProLeuArgSerLeuCysGlyArgAlaLeuValAlaLeuIlePh 17
288 TGCTCGCGGCTCTCGCGGATGGGAGAGAGAGAGAGAGAGAGAGAGAG 337
   |||||
17 eAlaCysGlyValAlaGlyValGlnSerGluGluArgGlyPheProPro 34
338 ACAGGCGCACT...CCGCTTTTGCACACCCAGAGATATGAGCCACCC 384
   |||||
34 laGlyAlaThrProProAlaLeuArgThrGlyGluIleValAlaProPro 50
385 ACTAAGACCTTATGGCCCAAGGCTTCCAGCCAGCTCTGGCGGGTGT 434
   |||||
51 ThrLysThrPheTrpProArgGlySerAsnAlaSerLeuProArgSer 67
435 GGCACCTCGCGAGGTGCTTAAGGAGACAGCGGCGGATCTCCGCCAC 484
   |||||
67 rSerProProGlnMetProLysGlyGlyArgMetAlaGlyProProAla 84
485 GCACCATCTCCCTCCCGCTGCCAGGAGCCCATCGAGATCAAGAGACT 534
   |||||
84 rgThrLeuThrProProCysGluGlyProIleGluIleLysAspThr 100
535 TTCAAATACATCAACACGGTTGTCTCTGCTTGTCTGCTGCTGGGAT 584
   |||||
101 PheLysThrIleAsnThrValValSerCysLeuValPheValLeuGly 117
585 CATCGGGAACCTCCACACTTCTGAGAAATATCTACAAGAACAAAGTGC 634
   |||||
117 eIleGlyAsnSerThrLeuLeuArgIleIleThrLysAsnLysCysMe 134
635 GAAACGGTCCCAATATCTTGCATCCCGCTGGCTCTGGGAGACCTCTG 684
   |||||
134 rgAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeu 150
685 CACATCGTCTTACATCCCTATCAATGTCTACAAGCTGCTGCAGAGGA 734
   |||||
151 HisIleIleIleAspIleProIleAsnValThrLysLeuLeuAlaGlu 167
735 CTGGCCATTGGAGCTGAGATGTGAAGCTGGTGGCTTTCATACAGAAAG 784
   |||||
167 pTrpProPheGlyValGluMetCysLysLeuValProPheIleGlnLys 184
785 CTCTCGTGGGAATCACTGCTGAGTCTATGCTCTGCTGAGTATTGACA 834
   |||||
184 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAsp 200
835 TATCGAGCTGCTCTCTTGGAGTAGAATTAAGGAATTCGGTTCACAA 884
   |||||
201 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValPro 217
885 ATGCAGCAGCAGTAGAATTTGTTGATTTGGGTGGTCTCTGTGTTCTGG 934
   |||||
217 sTrpThrAlaValAlaGluIleValLeuIleTrpValValSerValVal 234
935 CTCTCCCTGAAGCCATAGGTTTGTATATATATAGGATGGAGTACAAAG 984
   |||||
234 laValProGluAlaLeuGlyPheAspMetIleThrThrAspThrLys 250
985 AGTTATCTCGGAATCTGCTTCTCATCCCGTTCAGAGACAGCTTCAT 1034
   |||||
251 AsnArgLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPhe 267
1035 GCAGTTTTACAAGACAGCAAAAGATTGGTGGCTGTTCAGTTCTTATTCT 1084
   |||||
267 tGlnPheThrLysThrAlaLysAspTrpTrpLeuPheSerPheThrPhe 284
1085 GCTTGCCATTGGCCATCACTGCATTTTTTTATACACTAATGACCTGTGA 1134
   |||||
284 ysLeuProLeuAlaIleThrAlaPhePheThrLeuMetThrCysGlu 300
1135 ATCTTGAGAAGAAAGTGCATCCAGATTGCTTTAATCATCATCCTAAA 1184
   |||||

```

301 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLeu 317
 1185 GCAGAGCGGAAGTGGCAAAACCGCTTTTGGCTGGCTGCTGCTTTG 1234
 317 sGlnArgGluValAlaLysThrValPheCysLeuValLeuValPheA 334
 1235 CCCTCTGCTGCTTCCCTTCACCTCAGCAGGATTCGAAGCTCACTCTT 1284
 334 laLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 350
 1285 TATAATCAGATGATCCCACTAGATGTCACACTTTTGAGCTTCTGTTGGT 1334
 351 TyrAspGlnAsnAspSerAsnArgCysGluLeuLeuSerPheLeuLeu 367
 1335 ATTGAGCTATATTGGTATCAATGAGCTTCACTGAATTCCTGCATTAACC 1384
 367 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 384
 1385 CAATTGCTCTGTATTGGTGAGCAAAAGATTCAAAACTGCTTTAAGTCA 1434
 384 rolleAlaLeuTyLeuValSerLysArgPheLysAsnCysPheLysSer 400
 1435 TGCTTATGCTGCTGCCAGTCATTGAAGAAAAACAGTCTTGGAGGA 1484
 401 CysLeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGlu 417
 1485 AAAGCAGTCGTGCTTAAGTTCAAGCTTAATGATCAGGATATGACAACT 1534
 417 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsn 434
 1535 TCGCTTCCAGTAATAATACAGCTCACT 1563
 434 heArgSerSerAsnLysTySerSerSer 443

seq_name: SwissProt_40:ETBR_HORSE

seq_documentation_block:

ID ETBR_HORSE STANDARD; PRT; 443 AA.
 AC O62709; 077508;
 DT 15-DEC-1998 (Rel. 37, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
 DE Non-selective type).
 GN EDNRB.
 OS Equus caballus (Horse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
 OX NCBI_TaxID=9796;
 RN [1]
 RP SEQUENCE FROM N.A., AND VARIANT OLWS LYS-126.
 RC TISSUE=Liver;
 RX MEDLINE=98248431; PubMed=9580670;
 RA Yang G.C., Croaker D., Zhang A.L., Manglick P., Cartmill T., Cass D.;
 RT "A dinucleotide mutation in the endothelin-B receptor gene is
 RT associated with lethal white foal syndrome (LWFS); a horse variant of
 RT Hirschsprung disease".
 RL Hum. Mol. Genet. 7:1047-1052(1998).
 RN [2]
 RP SEQUENCE OF 8-443 FROM N.A., AND VARIANT OLWS LYS-126.
 RX MEDLINE=98191732; PubMed=9530628;
 RA Santschi E.M., Purdy A.K., Valberg S.J., Vrotsos P.D., Kaese H.,
 RA Mickelson J.R.;
 RT "Endothelin receptor B polymorphism associated with lethal white foal
 RT syndrome in horses";
 RL Mamm. Genome 9:306-309(1998).
 CC -1- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
 CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
 CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- DISEASE: DEFECTS IN EDNRB ARE A CAUSE OF OVERO LETHAL WHITE
 CC SYNDROME (OLWS) ALSO KNOWN AS LETHAL WHITE FOAL SYNDROME (LWFS).
 CC IT IS AN INHERITED SYNDROME OF FOALS BORN TO AMERICAN PAINT HORSE

CC PARENTS OF THE OVERO COAT-PATTERN LINEAGE. AFFECTED FOALS ARE
 CC TOTALLY OR ALMOST TOTALLY WHITE AND DIE WITHIN DAYS FROM
 CC COMPLICATIONS DUE TO INTestinal AGALIONOSIS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed, usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC

DR EMBL; AF019072; AAC25983.1; -;
 DR EMBL; AF038900; AAC23486.1; -;
 DR InterPro; IPR000276; GPCR_Rhodpsn.
 DR Pfam; PF00001; 7tm_1; 1.000000.
 DR PRINTS; PR00237; GPCRRHODPSN.
 DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR PROSITE; PS00262; G_PROTEIN_RECEP_F1_2; 1.
 DR KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Lipoprotein; Palmitate; Disease mutation.
 FT SIGNAL 1 26 POTENTIAL
 FT CHAIN 27 443 ENDOTHELIN B RECEPTOR.
 FT DOMAIN 27 102 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 103 127 1 (POTENTIAL).
 FT DOMAIN 128 138 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 139 164 2 (POTENTIAL).
 FT DOMAIN 165 176 3 (POTENTIAL).
 FT TRANSMEM 177 198 4 (POTENTIAL).
 FT DOMAIN 199 219 5 (POTENTIAL).
 FT TRANSMEM 220 244 6 (POTENTIAL).
 FT DOMAIN 245 272 7 (POTENTIAL).
 FT TRANSMEM 273 297 8 (POTENTIAL).
 FT DOMAIN 298 325 9 (POTENTIAL).
 FT TRANSMEM 326 351 10 (POTENTIAL).
 FT DOMAIN 352 363 11 (POTENTIAL).
 FT TRANSMEM 364 390 12 (POTENTIAL).
 FT DOMAIN 391 443 13 (POTENTIAL).
 FT CARBOHYD 60 60 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT DISULFID 175 256 BY SIMILARITY.
 FT LIPID 403 403 PALMITATE (POTENTIAL).
 FT LIPID 404 404 PALMITATE (POTENTIAL).
 FT LIPID 406 406 PALMITATE (POTENTIAL).
 FT VARIANT 126 126 I -> K (IN OLWS).
 FT CONFLICT 11 11 V -> A (IN REF. 2).
 FT CONFLICT 43 44 FE -> SG (IN REF. 2).
 FT CONFLICT 64 66 PRL -> QRS (IN REF. 2).
 FT CONFLICT 79 79 A -> E (IN REF. 2).
 FT CONFLICT 92 92 E -> D (IN REF. 2).
 FT CONFLICT 297 297 E -> M (IN REF. 2).
 FT CONFLICT 369 369 E -> D (IN REF. 2).
 FT CONFLICT 400 400 W -> S (IN REF. 2).
 SQ SEQUENCE 443 AA; 50006 MW; 74381DDFCF9A9AF4 CRC64;

alignment_scores:

Quality: 2085.50 Length: 443
 Ratio: 4.989 Gaps: 1
 Percent Similarity: 94.357 Percent Identity: 89.165

alignment_block:

US-09-931-157-2 x ETBR_HORSE

Align seg 1/1 to: ETBR_HORSE from: 1 to: 443

238 ATGCAGCGCCCTCCAAAGTCTGCGGACGCGCCCTGTTGCGCTGTTCT 287
 |||||
 1 MetGlnProLeuProThrLeuCysGlyArgValLeuValAlaLeuIle 17
 288 TGCCTGCGCGCTGTGCGGATCTGGGAGAGAGAGAGAGCTTCCCGCTG 337
 |||||
 17 ualaCysGlyValAlaGlyValGlnGlyGluArgPheProA 34

```
338 ACAGGCCACT...CCGTTTTTCAAAACCGCAGAGATAATGACGCCACCC 384
|||||
34 laArgAlaThrProProLeuLeuGlyPheGluGluIleMetThrProPro 50
|||||
385 ACTAGACCTTATGCGCCCAAGGGTCCAAACCCAGCTCTGGCGCGTCTGTT 434
|||||
51 ThrLysThrSerTrpProThrGlySerAsnAlaSerValProArgLeuSe 67
|||||
435 GCACCTCGCGAGGTGCCTAAAGGAGACAGGACGCGCAGGATCTCGGCCAC 484
|||||
67 rlaProProGlnMetProLysAlaGlyArgThrAlaGlyAlaGlnArgA 84
|||||
485 GCACCATCTCCCTCCCGCGTCCCAAGGACCCATCGAGATCAAGGAGACT 534
|||||
84 rgThrLeuProProProProCysGluArgThrIleGluIleLysGluThr 100
|||||
535 TTCAATACATCAACACGCTGTGTCCTGCTGCTGCTGCTGCTGCTGCTG 584
|||||
101 PheLysTyrIleAsnThrValValSerCysLeuValPheValIleGlyI 117
|||||
585 CATCGGGAATCCACACTTCTCAGAAATATCTACAAGAACAAAGTGCATGC 634
|||||
117 elleGlyAsnSerThrLeuLeuArgIleIleIleIleIleIleIleIle 134
|||||
635 GAAACGGTCCCAATATCTTGATCGCCAGCTTGCGTCTGGGAGACCTGCTG 684
|||||
134 rgAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 150
|||||
685 CACATCGTCATGACATCCCTCAATGCTACAAAGCTGCTGCGCAGGA 734
|||||
151 HisIleIleAspIleProIleAsnValTyrLysLeuLeuAlaGluAs 167
|||||
735 CTGGCATTGGAGTGAGATGTGAAGCTGGTGGCTTTTCATACAGAAAG 784
|||||
167 ptrProPheGlyValGluMetCysLysLeuValProPheIleGlnLysA 184
|||||
785 CTCCGTGGGAATCACTGTGCTGAGTCTATGTGCTCTGAGTATTGACAGA 834
|||||
184 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 200
|||||
835 TATCAGCGTGTCTCTTGGAGTAGAATTAAGGAATTTGGGGTTCCTCAA 884
|||||
201 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValPro 217
|||||
885 ATGCAGCAGTAGAAATTTGATTTGGTGGTGTCTCTGCGTTCTGG 934
|||||
217 strpThrAlaValGluIleValIleIleIleIleIleIleIleIleIle 234
|||||
935 CTGTCCTCGAAGCCATAGTTTGTATATTAATACGATGACACTACAAAGA 984
|||||
234 laValProGluAlaValGlyPheAspMetIleThrAlaAspTyrLysGly 250
|||||
985 AGTTATCTCGGATCTGCTGCTTCATCCGCTCAGAACAGACGCTTCAT 1034
|||||
251 SerTyrLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPhe 267
|||||
1035 GCAGTTTTTCAAGACAGCAAAAGATTGGTGGTGTTCAGTTCTATTCTT 1084
|||||
267 tGlnPheTyrLysAsnAlaLysAspTrpLeuPheSerPheTyrPheC 284
|||||
1085 GCTTGCCATTGGCCATCACTGCACTTTTATATACACTAATGACCTCGAA 1134
|||||
284 ysLeuProLeuAlaIleThrAlaPhePheTyrThrLeuGluThrCysGlu 300
|||||
1135 ATGTTGAGAAAGAAAGTGGCATGCGATGCTTTAAATGATCACCATAA 1184
|||||
301 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeu 317
|||||
1185 GCAGACACGGGAGTGGCCAAACCGCTCTTTGCTGGTCTGCTGCTG 1234
|||||
317 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuAlaPhe 334
|||||

seq_name: SwissProt_40:ETBR_RAT
seq_documentation_block:
ID ETBR_RAT STANDARD; PRT; 442 AA.
AC P21451;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
DE Non-selective type).
GN EDNRB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=92191882; PubMed=1312429;
RA Hori S., Komatsu Y., Shigemoto R., Mizuno N., Nakanishi S.;
RT "Distinct tissue distribution and cellular localization of two
RT messenger ribonucleic acids encoding different subtypes of rat
RT endothelin receptors."
RL Endocrinology 130:1885-1895(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=91080924; PubMed=2175397;
RA Sakurai T., Yanagisawa M., Takuwa Y., Miyazaki H., Kimura S.,
RA Goto K., Masaki T.;
RT "Cloning of a cDNA encoding a non-isopeptide-selective subtype of the
RT endothelin receptor."
RL Nature 348:732-735(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=93382424; PubMed=8371713;
RA Cheng H.F., Su Y.M., Yeh J.R., Chang K.J.;
RT "Alternative transcript of the nonselective-type endothelin receptor
RT from rat brain."
RL Mol. Pharmacol. 44:533-538(1993).
CC -!- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
```

CC -!- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN CELL TYPES OF A
 CC VARIETY OF TISSUES.
 CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; X57764; CAA40916.1; -;
 DR EMBL; S65355; AAB28172.1; -;
 DR PIR; S13425; S13425.
 DR PIR; A49197; A49197.
 DR GCRDb; GCR_0137; -;
 DR InterPro; IPR000276; GPCR_Rhodopsn.
 DR Pfam; PF00001; 7tm_1; 1.
 DR PRINTS; PR00237; GPCRHHODOPSN.
 DR PROSITE; PS00237; G-PROTEIN_RECEP_FL_1; 1.
 DR PROSITE; PS00262; G-PROTEIN_RECEP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Lipoprotein; Palmitate.
 FT SIGNAL 1 26
 FT CHAIN 27 442
 FT DOMAIN 27 101
 FT TRANSMEM 102 126
 FT DOMAIN 127 137
 FT TRANSMEM 138 163
 FT DOMAIN 164 175
 FT TRANSMEM 176 197
 FT DOMAIN 198 218
 FT TRANSMEM 219 243
 FT DOMAIN 244 271
 FT TRANSMEM 272 296
 FT DOMAIN 297 324
 FT TRANSMEM 325 350
 FT DOMAIN 351 362
 FT TRANSMEM 363 389
 FT DOMAIN 390 442
 FT DISULFID 174 255
 FT CARBOHYD 60 60
 FT CARBOHYD 97 97
 FT LIPID 402 402
 FT LIPID 403 403
 FT LIPID 405 405
 FT CONFLICT 86 69
 SQ SEQUENCE 442 AA; 49454 MW; BDA628D78E8071CB CRG64;

alignment_scores:
 Quality: 2070.00 Length: 443
 Ratio: 4.940 Gaps: 2
 Percent Similarity: 94.582 Percent Identity: 88.713
 alignment_block:
 US-09-931-157-2 x ETBR_RAT ..
 Align seg 1/1 to: ETBR_RAT from: 1 to: 442
 238 ATGACGCGCCTCAAGCTGTGGGACGCGCCTGTGGTGGTGTCT 287
 ||||| ||| ||||| ||||| ||||| ||||| ||||| |||||
 1 MetGlnSerSerAlaSerArgCysGlyArgAlaLeuValAlaLeuLeuLeu 17
 288 TGCGTCGGCCTGTGCGGATCTGGGAGAGAGAGAGCTTCCCGCTG 337
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 17 uAlaCysGlyLeuGlyValTrpGlyGlyLysArgGlyPheProProA 34
 338 ACAGGGCCACTCG...CTTTTGCACACCGCAGAGATAATGACGCCACC 384
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 34 laGlnAlaThrProSerLeuLeuGlyThrLysGluValMetThrProPro 50

385 ACTAAGACCTTATGCGCAAGGGTTCCACAGCCAGTCTGCGCGGTCT 434
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 51 ThrLysThrSerTrpThrArgGlySerAsnSerSerLeuMetArgSerSe 67
 435 GGCACCTCGGAGGTCCCTAAGGAGACAGGACGCGAGGATCTCGCCAC 484
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 67 rAlaProAlaGluValThrLysGlyGlyArgValAlaGlyValProProA 84
 485 GCACCATCTCCCTCCCGCGGCGCCAGGACCCATCGAGATCAAGGAGACT 534
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 84 rGSerPhe...ProProProCysGlnArgLysIleGluIleAsnLysThr 99
 535 TTCAATATACATCAACACGGTGTGTCTGCTGCTGCTGCTGCTGCTGCT 584
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 100 PheLysTyrIleAsnThrIleValSerCysLeuValPheValLeuGlyIle 116
 585 CATCGGGAACCTCCACACTTCTGAGAAATTTATACAAAGAACAAAGTCATGC 634
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 116 eileGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetA 133
 635 GAACGGTCCCAATATCTTGATCGCCAGCTGGCTGCTGGGAGACCTGCTG 684
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 133 rGAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeu 149
 685 CACATCGTCATTGACATCCCTATCAATGCTACAAAGCTGCTGGCAGAGGA 734
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 150 HisIleIleIleAspIleProIleAsnAlaTyrLysLeuLeuAlaGlyAs 166
 735 CTGGCCATTTGGAGCTGAGATGTGTAAAGCTGGTGGCTTTCATACAGAAG 784
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 166 pTrpProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysA 183
 785 CCTCGGCGGAATCAGCTGTGCTGAGTCTATGTGCTCTGCTGATTTACAGA 834
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 183 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 199
 835 TATCAGCTGTGCTTCTTGGAGTAGAATTAAGGAATTTGGGTTCCAAA 884
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 200 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProly 216
 885 ATGACAGCAGCTAGAAATTTGTTTGGTGGTCTCTGCTGCTGCTGCTG 934
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 216 sTrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuA 233
 935 CTGTCCTCGAAGCCATAGTGTGATATTAATACGATGCGATGACAAAGGA 984
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 233 laValProGluAlaIleGlyPheAspValIleThrSerAspTyrLysGly 249
 985 AGTTATCTGCGAATCTGCTGCTTCCATCCCGTTCAAGAGACAGCTTTTCAT 1034
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 250 LysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMe 266
 1035 GCAGTTTTTACAGACAGCAAGATTTGGTGGCTGTTCAGTTTCTATTCT 1084
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 266 tGlnPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheC 283
 1085 GCTTGCCATTTGCGCATCTGCTGCTATTTTATACACTAATACCTGTGAA 1134
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 283 ysLeuProLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGlu 299
 1135 ATGTTGAGAAAGAAAGTGGCATGCGATTCAGATTGCTTTAAATGATCACTAAA 1184
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 300 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 316
 1185 GCAGAGACGGGAAGTGGCCAAAACCGTCTTTTGGCTGCTCTGCTGCTTTG 1234
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 316 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 333
 1235 CCCTCTGCTGGCTTCCCTTCCCTTCCCTCAGCAGGATTTCTGAAGCTACTCT 1284
 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
 333 laLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 349
 1285 TATATCAGAAATGATGCCAATAGATGTCAACTTTTGTGCTTCTGCTGCT 1334


```

150 HisIleIleIleAspIleProIleAsnThrTyrLysLeuLeuAlaGluAs 166
735 CTGCCATTGGAGCTGAGATGTGTAAGCTGGTCCCTTTTCATACAGAAAG 784
|||||
166 pTrpProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysA 183
|||||
785 CCTCCGTGGGAATCACTGCTGCTGAGTCTATGTCTCTGAGTATTGACAGA 834
|||||
183 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 199
|||||
835 TATCGAGCTGTCTCTTCTGGAGTAGAATTAAGGAATTCGGGTCCAAA 884
|||||
200 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLy 216
|||||
885 ATGCACAGCAGTAGAATAATTTTGGTGGTCTCTGTGGTCTCTGG 934
|||||
216 sTrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuA 233
|||||
935 CTGTCCCTGAGCCATAGGTTTGTATATATACGATGGACTCAAGGA 984
|||||
233 laValProGluAlaIleGlyPheAspMetIleThrSerAspTyrLysGly 249
|||||
985 AGTTATCTCGAATCTGCTTCATCCCGTTTCAGAACACGCTTTCAT 1034
::: |||||:::|||||:::|||||
250 LysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaPheMe 266
|||||
1035 GCAGTTTTTCAACAGACAGAAAGATGGTGGCTTCAGTTCTATTTCT 1084
|||||
266 tGlnPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheC 283
|||||
1085 GCTTGCCATTGGCCATCACTGCATTTTATATACACTAATGACCTGTGAA 1134
|||||
283 ysLeuProLeuAlaIleThrAlaValPheTyrThrLeuMetThrCysGlu 299
|||||
1135 ATGTTGAGAAAGAAAGTGCATGCAGATGCTTTAAATGATCACCTAAA 1184
|||||
300 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 316
|||||
1185 GCAGACAGGGAGTGGCCAAACCGCTTTTTCGCTGGTCTGTCTTGTG 1234
|||||
316 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPhea 333
|||||
1235 CCCTCTGCTGGCTTCCCTTCACCTCAGCAGGATTCGAGCTCACTCTT 1284
|||||
333 laLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 349
|||||
1285 TATATCAAGATCATCCCAATAGATGTAACCTTTGAGCTTCTGTGGT 1334
|||||
350 TyrAspGlnSerAsnProHisArgCysGluLeuLeuSerPheLeuLeuVa 366
|||||
1335 ATTGGACTATATTGGTATCAACATGCTTCACCTGAATTCCTGCATTAAAC 1384
|||||
366 lleuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnP 383
|||||
1385 CAATTGCTCTGATTTGGTGAGCAAAAGATTCAAAACACTGCTTTAAGTCA 1434
|||||
383 roileAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSer 399
|||||
1435 TGCATTATGCTGTGGTCCAGTCATTTGAAGAAAACAGTCTCTGGAGGA 1484
|||||
400 CysLeuCysCysTrpCysGlnThrPheGluGluLysGlnSerLeuGluG 416
|||||
1485 AAAGCAGTCTGCTTAACTTCAAGCTAATGATCAGCATATGACAACT 1534
|||||
416 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnP 433
|||||
1535 TCCGTTCCAGTAAATAATACAGCTCATCT 1563
|||||
433 heArgSerSerAsnLysTyrSerSerSer 442
|||||
seq_name: SwissProt_40:ETBR_BOVIN
seq_documentation_block:

```

```

ID ETBR_BOVIN STANDARD; PRT; 441 AA.
AC P28088;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Endothelin B receptor precursor (ET-B) (Endothelin receptor
DE Non-selective type).
GN EDNRB.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=92078223; PubMed=1660473;
RA Saito Y., Mizuno T., Itakura M., Suzuki Y., Ito T., Hagiwara H.,
RA Hirose S.;
RT "Primary structure of bovine endothelin ETB receptor and
RT identification of signal peptidase and metal proteinase cleavage
RT sites.";
RL J. Biol. Chem. 266:23433-23437(1991).
RN [2]
RP SEQUENCE OF 124-127; 262-269; 304-315; 417-421 AND 424-432.
RC TISSUE=Lung;
RX MEDLINE=91358493; PubMed=1653249;
RA Kozuka M., Ito T., Hirose S., Iodhi K.M., Hagiwara H.;
RT "Purification and characterization of bovine lung endothelin
RT receptor.";
RL J. Biol. Chem. 266:16892-16896(1991).
RN [3]
RP POST-TRANSLATIONAL MODIFICATIONS.
RC TISSUE=Lung;
RX MEDLINE=98086265; PubMed=9422751;
RA Roos M., Soskic V., Poznanovic S., Godovac-zimmermann J.;
RT "Post-translational modifications of endothelin receptor B from
RT bovine lungs analyzed by mass spectrometry.";
RL J. Biol. Chem. 273:924-931(1998).
CC -!- FUNCTION: NON-SPECIFIC RECEPTOR FOR ENDOTHELIN 1, 2, AND 3.
CC MEDIATES ITS ACTION BY ASSOCIATION WITH G PROTEINS THAT ACTIVATE
CC A PHOSPHATIDYLINOSITOL-CALCIUM SECOND MESSENGER SYSTEM.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; D10994; BAA01762.1;
CC EMBL; D10989; BAA01762.1; JOINED.
CC EMBL; D10990; BAA01762.1; JOINED.
CC EMBL; D10991; BAA01762.1; JOINED.
CC EMBL; D10992; BAA01762.1; JOINED.
CC EMBL; D10993; BAA01762.1; JOINED.
CC EMBL; D90456; BAA14422.1;
CC PIR; A41591; A41591.
CC GCRDB; GCR_0237;
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEPTOR_F1_1; 1.
CC PROSITE; PS00262; G_PROTEIN_RECEPTOR_F1_2; 1.
CC G-protein coupled receptor; Transmembrane; Signal; Lipoprotein;
KW Palmitate; Phosphorylation.
FT SIGNAL 1
FT CHAIN 26
FT DOMAIN 27 441 ENDOTHELIN B RECEPTOR.
FT DOMAIN 27 100 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 101 125 1 (POTENTIAL).
FT DOMAIN 126 136 CYTOPLASMIC (POTENTIAL).

```

```
FT TRANSEM 137 162 2 (POTENTIAL).
FT DOMAIN 163 174 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 175 196 3 (POTENTIAL).
FT DOMAIN 197 217 4 (POTENTIAL).
FT TRANSEM 218 242 5 (POTENTIAL).
FT DOMAIN 243 270 EXTRACELLULAR (POTENTIAL).
FT TRANSEM 271 295 6 (POTENTIAL).
FT DOMAIN 296 323 7 (POTENTIAL).
FT TRANSEM 324 349 8 (POTENTIAL).
FT DOMAIN 350 361 9 (POTENTIAL).
FT TRANSEM 362 388 10 (POTENTIAL).
FT DOMAIN 389 441 11 (POTENTIAL).
FT DISULFID 173 234 BY SIMILARITY.
FT LIPID 402 402 PALMITATE.
FT LIPID 404 404 PALMITATE.
FT MOD_RES 304 304 PHOSPHORYLATION.
FT MOD_RES 418 418 PHOSPHORYLATION.
FT MOD_RES 434 434 PHOSPHORYLATION (OR 435).
FT MOD_RES 438 438 PHOSPHORYLATION.
FT MOD_RES 439 439 PHOSPHORYLATION.
FT MOD_RES 440 440 PHOSPHORYLATION.
FT MOD_RES 441 441 PHOSPHORYLATION.
SQ SEQUENCE 441 AA: 49371 MW: 634462C544DBD2 CRC64;

alignment_scores:
  Quality: 2046.50 Length: 443
  Ratio: 4.919 Gaps: 2
Percent Similarity: 93.905 Percent Identity: 88.488

alignment_block:
US-09-931-157-2 x ETBR_BOVIN ..
Align seg 1/1 to: ETBR_BOVIN from: 1 to: 441

238 ATGGAGCGCCCTCAAGTCTGTGGGAGCGCCCTGTGGCGGTCT 287
|||||
1 MetGlnProLeuProSerLeuCysGlyArgAlaLeuValAlaLeuLe 17
288 TGCTGCGCGCTGTCGGAGGAGGAGGAGGAGGCTTCCGCGCTG 337
|||||
17 uAlaCysGlyValAlaGlyIleGlnAlaGluGluArgGluPhePro 34
338 ACAGGGCCACT...CCGCTTTGCAACCGCAGAGATAATGACGCC 384
|||||
34 laGlyAlaThrGlnProLeuProGlyThrGlyGluMetMetGluThr 50
385 ACTAAGACCTTATGCCCCAAGGTTGCCAACCCAGCTCTGGCGG 434
|||||
51 ThrGluThrSerTrpProGlyArgSerAsnAlaSerAspProArgSe 67
435 GGCACCTCGGAGGTGCTTAAGGAGACAGGAGCGGATCTCCGCC 484
|||||
67 rAlaThrProGlnIleProArgGlyGlyArgMetAlaGlyIlePro 84
485 GCACACTCTCCCTCCCGGCGGAGGAGGAGGAGGAGGAGGAG 534
|||||
84 rgThr.....ProProCysAspGlyProIleGluIleLysGluThr 98
535 TTCAAATACATCAACACCGTTGTCTCTGCTGCTGCTGCTGGG 584
|||||
99 PheLysTyriIleAsnThrValValSerCysLeuValPheValL 115
585 CATCGGGAGCTCCACACTCTTGAGAAATATCTACAAGAACAGT 634
|||||
115 eIleGlyAsnSerThrLeuLeuArgIleIleIleIleIleIleI 132
635 GAAACGGTCCCAATATCTTGATCGGAGCTTGGCTCTGGGAGC 684
|||||
132 rGAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAsp 148
685 CACATCGCATTCATCCCTATCAATGTCTACAAGCTGCTGGCAG 734
|||||
```

seq_name: SwissProt_40:ETBR_COTJA

seq_documentation_block:

```
149 HisIleIleIleAspIleProIleAsnThrTyrLysLeuLeuAlaLysAs 165
735 CTGGCCATTTGGAGCTGAGATGTCTAAGCTGGTGCCTTTTCATACAGAAG 784
|||||
165 pTrpProPheGlyValGluMetCysLysLeuValProPheIleGlnLysA 182
785 CCTCCGTTGGGAATCACTGTGCTGAGTCTATGTCTGCTGAGTATTGACAGA 834
|||||
182 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 198
835 TATCAGAGCTGTCTCTTGGAGTAGAATTAAGGAATTTGGGTTCCTCAA 884
|||||
199 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLy 215
885 ATGGACAGCAGTAAATTTGTTGATTGGTGGTCTCTCTGTGGTTCGG 934
|||||
215 sTrpThrAlaValGluIleValLeuIleTrpValValSerValValLeuA 232
935 CTGTCCTCGAAGCCATAGGTTTGTATATATACGATGAGTACGATACAGA 984
|||||
232 laValProGluAlaValGlyPheAspIleIleThrSerAspHisIleGly 248
985 AGTTATCTGCGAATCTGCTTCTCATCCCTTCAGAGACAGACAGCTTTCA 1034
|||||
249 AsnLysLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMe 265
1035 GCAGTTTTACAGACAGCAAAAGATTGGTGGTCTGTCTCACTTCTATTCT 1084
|||||
265 tGlnPheTyrLysThrAlaLysAspTrpTrpLeuPheSerPheTyrPheC 282
1085 GCTTGCATTCGCTCATCTGCTATTTTATACACTAATGACCTGTGAA 1134
|||||
282 ysLeuProLeuAlaIleThrAlaLeuPheTyrThrLeuMetThrCysGlu 298
1135 ATGTTGAGAAAGAAAGTGGCATGCAGATGCTTTTAAATGATCAGCTAAA 1184
|||||
299 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 315
1185 GCAGAGACGGGAAGTGGCCAAACCGCTTTTGGCTGCTGCTGCTGCTTTG 1234
|||||
315 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 332
1235 CCCTCTGCTGCTTCCCTTCCCTCAGCTCAGCAGGATTCGAGCTCACTCTT 1284
|||||
332 laLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 348
1285 TATATCAGAAATGATCCCAATAGATGTCAACTTTTGGAGCTTCTCTGCT 1334
|||||
349 TyrAspGlnHisAspProArgArgCysGluPheLeuSerPheLeuLeuVa 365
1335 ATTGGACTATATTGGTATCAACATGGCTTCACTGAATTCCTGCAATTAAC 1384
|||||
365 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 382
1385 CAATTGCTCTGTATTGGTGAGCAAAAGATTCAAAAAGCTGCTTTAAGTCA 1434
|||||
382 roIleAlaLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSer 398
1435 TGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1484
|||||
399 CysLeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGl 415
1485 AAAGCAGTCTGCTTAAAGTTCAAAGCTTAATGATCAGCGGATATGACAAC 1534
|||||
415 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsn 432
1535 TCCGTTCCAGTAATAATACAGCTCATCT 1563
|||||
432 heArgSerSerAsnLysTyrSerSerSer 441
seq_name: SwissProt_40:ETBR_COTJA
seq_documentation_block:
```



```

1420 AACTGCTTTAAGTATGCTTATGCTGGTGGCCAGTCATTTGAAGAAA 1469
|||||
301 AsnCysPheLysSerCysLeuCysCysTrpCysGlnSer...LysAspLe 316
1470 ACAGTCCCTGGAGGAAAGCAGCTGCTTAAAGTTCAAGCTAATGATC 1519
|||||
316 uLeuSerLeuGluArgGlnSerCysLeuLysPheLysAlaAsnAspH 333
1520 ACGGATATGACAACTCCGTTCCAGTAAATAATACAGCTCATCT 1563
|||||
333 IsGlyTyrAspAsnPheArgSerSerAsnLysTyrSerSerSer 347
seq_name: SwissProt_40:ETLR_RAT

seq_documentation_block:
ID ETLR_RAT STANDARD; PRT; 426 AA.
AC P26684;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Endothelin-1 receptor precursor (ET-A).
GN EDNR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=92191882; PubMed=1312429;
RA Hori S., Komatsu Y., Shigemoto R., Mizuno N., Nakanishi S.;
RT "Distinct tissue distribution and cellular localization of two
RT messenger ribonucleic acids encoding different subtypes of rat
RT endothelin receptors.";
RL Endocrinology 130:1885-1895(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=91195317; PubMed=1849646;
RA Lin H.Y., Kaji E.H., Winkel G.K., Ives H.E., Lodish H.F.;
RT "Cloning and functional expression of a vascular smooth muscle
RT endothelin 1 receptor.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:3185-3189(1991).
CC -1- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN VASCULAR SMOOTH
CC MUSCLE CELLS OF A VARIETY OF ISSUES, BRONCHIAL SMOOTH MUSCLE
CC CELLS, MYOCARDIUM, AND THE PITUITARY GLAND.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M60786; AAA41114.1; -
CC PIR; A40440; A40440.
CC PIR; B49197; B49197.
CC GCRDB; GCR_0136; -
CC InterPro; IPR000276; GPCR_Rhodpsn.
CC Pfam; PF00001; 7tm_1; 1.
CC PRINTS; PR00237; GPCRHHODOPSN.
CC PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
CC PROSITE; PS0262; G_PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
CHAIN 18 426 ENDOTHELIN-1 RECEPTOR.

```

```

FT DOMAIN 18 80 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 81 102 1 (POTENTIAL).
FT TRANSNEM 103 112 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 113 132 2 (POTENTIAL).
FT TRANSNEM 133 159 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 160 181 3 (POTENTIAL).
FT TRANSNEM 182 205 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 206 229 4 (POTENTIAL).
FT TRANSNEM 230 256 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 257 278 5 (POTENTIAL).
FT TRANSNEM 279 306 CYTOPLASMIC (POTENTIAL).
FT TRANSNEM 307 328 6 (POTENTIAL).
FT TRANSNEM 329 347 EXTRACELLULAR (POTENTIAL).
FT TRANSNEM 348 372 7 (POTENTIAL).
FT TRANSNEM 373 426 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 62 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 158 239 BY SIMILARITY.
FT CONFLICT 46 53 N -> D (IN REF. 2).
FT CONFLICT 53 53 Q -> R (IN REF. 2).
SQ SEQUENCE 426 AA; 48244 MW; E8A24683DBB78BB8 CRC64;

```

alignment_scores:

Quality: 1246.50 Length: 376
Ratio: 3.970 Caps: 5
Percent Similarity: 83.511 Percent Identity: 61.436

alignment_block:

```

US-09-931-157-2 x ETLR_RAT ..
Align seg 1/1 to: ETLR_RAT from: 1 to: 426
487 ACCATCTCCCTCCCTCCCG..... 504
|||||
51 ThrLeuGlnProProAsnLeuAlaLeuProSerAsnGlySerMethisgl 67
505 ....TGCCAAGGAGCCATCGAGATCAAGGAGACTTTCAAAATACATCAACA 550
|||||
67 TyrCysProGlnGlnThrLysIleThrAlaPheLysTyrIleAsnT 84
551 CGGTGTGTCTCCCTGCTGGTGGTGGGATCATCGGGAACGCCACA 600
|||||
84 hrValIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThr 100
601 CTTCTGAGAATTATACAAAGAACAGTCATCGGAAACGGTCCCAATAT 650
|||||
101 LeuLeuArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAl 117
651 CTTGATCGCCAGCTTGGCTCTGGGAGACCTGCTGCACATCGTCATTGACA 700
|||||
117 aLeuIleAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspL 134
701 TCCTTATCAATCTTACAGCTGCTGGCAGAGGAGCTGGCCA..... 741
|||||
134 euProIleAsnValPheLysLeuAlaGlyArgTrpPheAspHis 150
742 .....TTGGAGCTGAGATGTGTAAGCTGGTCCCTTCATACAGAAAGC 785
|||||
151 AsnAspPheGlyValPheLeuLysLeuPheProPheLeuGlnLysSe 167
786 CTCCGTGGGAATCATCTGCTGAGTCTATGCTCTGAGTATTGACAGAT 835
|||||
167 rSerValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgT 184
836 ATCGAGCTGCTGCTTCTTGGAGTACAGATTAAAGGAATGGGCTCCAAA 885
|||||
184 yrArgAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeu 200
886 TGCACAGCAGTAGAAATTTGTTGATTGGTGGTCTCTGTGGTCTTCGC 935
|||||
201 IleThrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAl 217

```


CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC -----

```

DR EMBL: D90348; BAA14359.1; -
DR EMBL: L06622; AA58447.1; -
DR EMBL: X61950; CAA43953.1; -
DR EMBL: S67127; AAB20407.1; -
DR EMBL: S63938; AAB20278.1; -
DR EMBL: S57498; CAB31573.1; ALT_SEQ.
DR EMBL: D11151; BAA01920.1; -
DR EMBL: D11145; BAA01920.1; JOINED.
DR EMBL: D11146; BAA01920.1; JOINED.
DR EMBL: D11147; BAA01920.1; JOINED.
DR EMBL: D11148; BAA01920.1; JOINED.
DR EMBL: D11149; BAA01920.1; JOINED.
DR EMBL: D11150; BAA01920.1; JOINED.
DR EMBL: S45956; AAB23644.1; -
DR EMBL: S81539; AAB36325.1; -
DR EMBL: S55772; AAB25212.1; -
DR PIR: JS0607; JS0607.
DR PIR: JT0972; JT0972.
DR PIR: S16980; S16980.
DR PIR: A44158; A44158.
DR PIR: A45481; A45481.
DR GCRDb: GCR_0403; -
DR GCRDb: GCR_0473; -
DR GCRDb: GCR_0491; -
DR GCRDb: GCR_0512; -
DR GCRDb: GCR_0602; -
DR GCRDb: GCR_0727; -
DR GCRDb: GCR_0737; -
DR GCRDb: GCR_1922; -
DR MIM: 131243; -
DR InterPro: IPR000276; GPCR_Rhodopsn.
DR Pfam: PF00001; 7tm1_1; 1.
DR PRINTS: PR00237; GPCR_HODOPSIN.
DR PROSITE: PS00237; G_PROTEIN_RECEPT_F1_1; 1.
DR PROSITE: PS0262; G_PROTEIN_RECEPT_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 17 POTENTIAL.
FT CHAIN 18 427 ENDOTHELIN-1 RECEPTOR.
FT DOMAIN 18 80 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 81 102 1 (POTENTIAL).
FT DOMAIN 103 112 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 113 132 2 (POTENTIAL).
FT DOMAIN 133 159 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 160 181 3 (POTENTIAL).
FT DOMAIN 182 205 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 206 229 4 (POTENTIAL).
FT DOMAIN 230 256 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 257 278 5 (POTENTIAL).
FT DOMAIN 279 306 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 307 328 6 (POTENTIAL).
FT DOMAIN 329 347 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 348 372 7 (POTENTIAL).
FT DOMAIN 373 427 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULFID 62 62 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CONFLICT 158 239 BY SIMILARITY.
FT CONFLICT 110 110 C -> Y (IN REF. 7).
FT CONFLICT 207 207 V -> A (IN REF. 5).
FT CONFLICT 234 234 E -> G (IN REF. 5).
FT CONFLICT 280 280 C -> G (IN REF. 5).
SQ SEQUENCE 427 AA; 48722 MW; 207272DA231404F CRC64;

```

alignment_scores:
 Quality: 1240.00 Length: 359
 Ratio: 3.987 Gaps: 5
 Percent Similarity: 86.630 Percent Identity: 63.510

alignment_block:

US-09-931-157-2 x ETIR_HUMAN ..

Align seg 1/1 to: ETIR_HUMAN from: 1 to: 427

```

505 TGCAAAGGACCCATCGAGATCAAGAGACATTCACAAATACATCAACACGGT 554
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
69 CysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnThrVa 85
555 TGTGTCTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 604
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
85 ILeSerCysThrIlePheIleValGlyMetValGlyAsnAlaThrLeu 102
605 TGAGAATTATCTACAAGAACAAAGTCATGCGAAGCGTCCCAATATCTTG 654
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
102 euArgIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeu 118
655 ATCGCCAGCTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
119 ILeAlaSerLeuAlaLeuGlyAspLeuIleTyrValValIleAspLeuPr 135
705 TATCAATGTCTACAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 741
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
135 oIleAsnValPheLysLeuAlaGlyArgTyrProPheAspHisAsnA 152
742 ..TTTGGAGCTGAGATGTGTAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 789
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
152 spPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSerSer 168
790 GTGGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 839
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
169 ValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTyrAr 185
840 AGCTGTTGCTCTTGTGAGTAGAATTAAAGGAATTGGGTCCCAAAATGGA 889
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
185 gAlaValAlaSerTyrPheArgValGlnGlyIleGlyIleProLeuVal 202
890 CAGCAGTAGAAATTTGTTGATTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
202 hrAlaIleGluIleValSerIleTyrIleLeuSerPheIleLeuAlaIle 218
940 CCTGAAGCATAGTTTGTATATAATACATGGACTACAAAGGAAGTTA 989
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
219 ProGluAlaIleGlyPheValMetValProPheGluTyrArgGlyGlu 235
990 TCTGCCAATCTGCTTCTCATCCCTTCAGAGACAGCTTTTCATCAGT 1039
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
235 nHisLysThrCysMetLeuAsnAlaThrSerLys.....PheMetGluP 250
1040 TTTACAAGACAGCAAAAGATTGGTGGCTGTTCAAGTTTCTATTCTGCTTG 1089
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
250 heTyrGlnAspValLysAspTyrTyrPheGlyPheTyrPheCysMet 266
1090 CCATTGGCCATCACTGCATTTTATACACTAATGACCTGTGAATGTT 1139
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
267 ProLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLe 283
1140 GAGAAAGAAAGTGGC...ATGCAGATTGCTTTAAATGATCACCCTAAGC 1186
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
283 uAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLys 300
1187 AGAGACGGGAAGTGGCAAAACCGTCTTTTGGCTGGTCTGCTGCTGCTGCTGCT 1236
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
300 InArgArgGluValAlaLysThrValPheCysLeuValValIlePheAla 316
1237 CTCTGCTGCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTTCCCTT 1286
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
317 LeuCysTyrPheProLeuHisLeuSerArgIleLeuLysLysThrValty 333
1287 TAATCAAGATGATCCCAATAGATGTGAACCTTTTGAGCTTCTCTGCTGAT 1336
    ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
333 rAsnGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeu 350

```

```
1337 TGGACTATATGGTATCAACAGCTGCTTCACGTGAATCTCTGCATTAACCCA 1386
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
350 etAspTyrlleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnPro 366
1387 ATTGCTCTGTATTGGTGACCAAGAGATTCAAAAGCTGCTTTAAGTCATG 1436
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
367 IleAlaLeuTyrrPheValSerLysLysPheLysAsnCysPheGlnSerC 383
1437 CTTATGCTGCTGGTGC...CAGTCATTGTAAGAAAAACAGCTCTGGAGG 1483
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
383 sLeuCysCysCysTyrrGlnSerLysSerLeuMetThrSerValProM 400
1484 AAAAGCAGCTGCTGCTTAAAGCTCAAGCTTAATGAT...CAGGATATGAC 1530
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
400 etAsnGlyThrSerIleGlnTriPlysAsnHisAspGlnAsnAsnHisAsn 416
1531 AACTTCGCTCCAGTAATAATACAGC 1557
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
417 ThrAspArgSerSerHisLysAspSer 425
seq_name: SwissProt_40:ETIR_BOVIN

seq_documentation_block:
ID ETIR_BOVIN STANDARD; PRT; 427 AA.
AC P21450;
DT 01-MAY-1991 (Rel. 18, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Endothelin-1 receptor precursor (ET-A).
GN EDNRA.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=91080923; PubMed=2175396;
RA Arii H., Hori S., Aramori I., Ohkubo H., Nakanishi S.;
RT "Cloning and expression of a cDNA encoding an endothelin receptor.";
RL Nature 348:730-732(1990).
CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X57765; CAA40917.1; -
DR FICR; S13424; S13424.
DR GCRdb; GCR_0005; -
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm_1; 1.
DR PRINTS; PR00237; GPCRHHODPSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_FL_1; 1.
DR PROSITE; PS00262; G_PROTEIN_RECEP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL 1 17
FT CHAIN 18 427
FT DOMAIN 18 80
FT TRANSMEM 81 102
FT DOMAIN 103 112
FT TRANSMEM 113 132
```

```
FT DOMAIN 133 159 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 160 181
FT DOMAIN 182 205 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 206 229
FT DOMAIN 230 256 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 257 278
FT DOMAIN 279 306 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 307 328
FT DOMAIN 329 347 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 348 372
FT DOMAIN 373 427 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 29 29 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 62 62 N-LINKED (GLCNAC...) (POTENTIAL).
FT DISULFID 158 239 BY SIMILARITY.
SQ SEQUENCE 427 AA; 48515 MW; 73CF4404CD9AC14 CRC64;

alignment_scores:
  Quality: 1238.00      Length: 359
  Ratio: 3.981         Gaps: 5
  Percent Similarity: 86.630 Percent Identity: 63.510

alignment_block:
US-09-931-157-2 x ETIR_BOVIN ..
Align seg 1/1 to: ETIR_BOVIN from: 1 to: 427

505 TGCCAAGGACCCATCGAGATCAAGGAGACTTTCRAATACATCAACACGGT 554
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
69 CysProGlnGlnThrLysIleThrSerAlaPheLysTyrlleAsnThrVa 85

555 TGTCTCTGCTGCTGTGTGCTGGGATCATCGGGAACCTCCACACATC 604
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
85 lIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThrLeuL 102

605 TGAGAAATATCTCAAGAACAAAGTCATCGGAAACGGTCCCAATATCTTG 654
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
102 euArgIleIleTyrrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeu 118

655 ATCGCCAGCTGCTGCTGGGAGACCTGCTGCACATCGTCATTGACATCCC 704
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
119 lIleAlaSerLeuAlaLeuGlyAspLeuIleTyrrValIleAspLeuPr 135

705 TATCAATGCTACAAAGCTGCTGGCAGAGGACTGGCCA..... 741
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
135 oIleAsnValPheLysLeuLeuAlaGlyArgTrpProPheGluGlnAsnA 152

742 ..TTTGGAGCTGAGATGTGAAGTGGTGGCTTTCATACAGAAAGCCTCC 789
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
152 spPheGlyValPheLeuCysLysLeuPheProPheLeuGlnLysSerSer 168

790 GTGGGAATCACTGCTGCAGTCTATGCTGCTGAGTATTGACAGATATCG 839
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
169 ValGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTyrr 185

840 AGCTGTGCTCTCTGGAGTAGAATTAAGGAATGGGGTTCACAAATGGA 889
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
185 gAlaValAlaSerTrpSerArgValGlnGlyIleGlyIleProLeuValT 202

890 CAGCAGTAGAAATTTGTTTGAATTTGGGGTCTCTGTTGTTCTGCTGTC 939
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
202 hrAlaIleGluIleValSerIleThrPheLeuSerPheIleLeuAlaIle 218

940 CCTGAAGCCATAGCTTTTGATATAATACGATGGACTACAAAGGAGTTA 989
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
219 ProGluAlaIleGlyPheValIleNetValProPheGluTyrrLysGlyAlaG 235

990 TCTCGGAATCTGCTTGTTCATCCGTTCAAGAACAGACGCTTCATGCAGT 1039
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
235 nHisArgThrCysMetLeuAsnAlaThrSerLys.....PheMetGluP 250

1040 TTTACAAGACAGCAAAAGATTGGTGGCTGTTCAGTTCTATTCTCTGCTTG 1089
      |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

250 heTyrGlnAspValLysAspTrpIrrLeuPheGlyPheTyrPheCysMet 266

11090 CCATTGGCCATCACTGCATTTTTTTTATACACTAATGACCTGTCAAATGTT 1139
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
267 ProLeuValcysThrAlaIlePheTyrThrLeuMetThrCysGluMetLe 283

11140 GAGAAAGAAAGTGGC...ATGCAGATTGCTTTTAATATCATCACTAAAGC 1186
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
283 uAsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysG 300

11187 AGACACGGGAAGTGCCAAAACCGTCCTTTTGCCTGGGCCCTGCTTTGGCC 1236
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
300 InArgArgGluValAlaLysThrValPheCysLeuValValIlePheAla 316

1237 CTCTGCTGGCTTCCCCTTCACTCAGCAGAGATTCTGAAGCTCACTCTTTA 1286
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
317 LeuCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTy 333

1287 TAATCAGAATGATCCCCAATAGATGTGAACCTTTTGAGCTTTTCTCTGTGAT 1336
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
333 rAspGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuLeuM 350

1337 TGGACTATATTGGTATCAACATGGCTTCATCGATTCTCTGCATTAAACCCA 1386
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
350 etAspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnPro 366

1387 ATTCTCTGTATTGTGGTACCAAGAGATTCAAAAACGCTTTTAAGTCATG 1436
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
367 IleAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCy 383

1437 CTTATGCTGCTGGTGGC...CAGTCATTTTGAAGAAAAACAGTCCTTGGAGG 1483
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
383 sLeuCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValProm 400

1484 AAAACGACGTCTGCTTTAAAGTTCAAAGCTAATGAT...CACGGGATATGAC 1530
|||||::: |||||::: |||||::: |||||::: |||||::: |||||:::
400 etAsnGlnTyrSerIleGlnTrpLysAsnHisGluGlnAsnAsnHisAsn 416

1531 AACTCCGTTCCAGTAATAAATACAGC 1557

417 ThrGluArgSerSerHisLysAspSer 425

seq name: SwissProt 40:ET1R_PIG

```
seq_documentation_block:
ID   ET18 PIG STANDARD:
PRT: 427 AA:
```

Q23010;
15-JUL-1998 (Re) 36 Created)

15-JUL-1998 (Rel. 36, Last sequence update)
15-JUL-1998 (Rel. 36, Last annotation update)

Endothelin-1 receptor precursor (ET-A).
FDNRA

Sus scrofa (Plg).

Mammalia; Eutheria; Cetartiodactyla; Suina; NCBI TaxID=9823.

[1] SEQUENCE FROM N A

TISSUE=Lung;
MEYTIME=0503105A. DUBMED=0504103.

Nishimura J., Aoki H., Chen X., Shikasho T.,
Yasuda H.

"Evidence for the presence of endothelin ETA

Br. J. Pharmacol. 115:1369-1376(1995).

ASSOCIATION WITH G PROTEINS THAT ACTIVATES CALCIUM SECOND MESSENGER SYSTEM THE BANKING

AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.

-!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROT

This SWISS-PROT entry is copyright. It is pro

```
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; S80652; AAB36014.1; .
DR GCRDb; GCR_1954; .
DR InterPro; IPR000276; GPCR_Rhodpsn.
DR Pfam; PF00001; 7tm.1; 1.
DR PRINTS; PR00237; GPCRRHODPSN
DR PROSITE; PS00237; G-PROTEIN_RECP_FL_1; 1.
DR PROSITE; PS0262; G-PROTEIN_RECP_FL_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT SIGNAL      1    17     POTENTIAL.
FT CHAIN       18   427    ENDOTHELIN-1 RECEPTOR.
FT DOMAIN      18    80     EXTRACELLULAR (POTENTIAL).
FT TRANSEM     81   102     1 (POTENTIAL).
FT DOMAIN     103   112     CYTOPLASMIC (POTENTIAL).
FT TRANSEM    113   132     2 (POTENTIAL).
FT DOMAIN     133   159     EXTRACELLULAR (POTENTIAL).
FT TRANSEM    160   181     3 (POTENTIAL).
FT DOMAIN     182   205     CYTOPLASMIC (POTENTIAL).
FT TRANSEM    206   229     4 (POTENTIAL).
FT DOMAIN     230   256     EXTRACELLULAR (POTENTIAL).
FT TRANSEM    257   278     5 (POTENTIAL).
FT DOMAIN     279   306     CYTOPLASMIC (POTENTIAL).
FT TRANSEM    307   328     6 (POTENTIAL).
FT DOMAIN     329   347     EXTRACELLULAR (POTENTIAL).
FT TRANSEM    348   372     7 (POTENTIAL).
FT DOMAIN     373   427     CYTOPLASMIC (POTENTIAL).
FT CARBOHYD    29    29     N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD    62    62     N-LINKED (GLCNAC. .) (POTENTIAL).
FT DISULPID    158   239     BY SIMILARITY.
SQ SEQUENCE    427 AA; 48694 MW; F09366051552D06D CRC64;
```

```
alignment_scores:
Quality: 1233.00          Length:.. 359
Ratio: 3.965              Gaps: 5
Percent Similarity: 86.630 Percent Identity: 63.231

alignment_block:
US-09-931-157-2 x ETIR_PIG ..

Align seg 1/1 to: ETIR_PIG from: 1 to: 427
```

```
505 TCCCAAGGACCATCGAGATCAAGGAGACTTTCAAATACATCAACACGGT 554
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
69 CysProGlnThrLysIleThrSerAlaPheLysTyrrIleAsnThrVa 85

555 TGTGCTCGCCTTGTTGCTGCTGGGATCATCGGAACCTCCACACTTC 604
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
85 lileSerCysThrIlePheIlleValcglyMetValGlyAsnAlaThrLeul 102

605 TCAGAATTATCTACAAGACAAGTCATCGCAAACGGTCCCATAATCTTG 654
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
102 euargillelletyrlGlnasnLyScysMetArgasnclpyroAsnalaleu 118

655 ATCGGCAGCTTGCTGCTGGGAGACCTGCTGCACATCGTCATTGACATCC 704
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
119 lleAlaserLeualauleuglyspleulletyrvalvalilleasePr 135

705 TTATCAATGTCTACAAGCTGCTGGCAGAGACTGGCCA..... 741
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
135 oileasnlValPheLysLeulaalaglyArgrtprProPheGluaAsnHisA 152

742 ...TTTGGAGCTGAGATGTGTAGCTGGTGCCTTTTCATACAGAAGCCGCC 789
||||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
152 spPhelgylValPheLeucystylsLeuPheProPheGlnLysSer 168

790 GTGGGAATCACGTGCTGAGTGTATGTGCTGTGAGTATTGACAGATATCG 839
```

250	heYrGlnAspValylsAspTrpTrpLeuPheGlyPheTyPheCysMet	266
1090	CCATTGGCCATCACTGCATTTTTTTATACACTAATGACCTGTGAATGTT	1139
267	ProLeuValCysThrAlaIlePheTyTrpLeuMetThrCysGluMetLe	283
1140	GAGAAAGAAAGTGGC...ATGCAGATTGCTTTAAATGATCACCTAAAGC	1186
283	uAsnArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysG	300
1187	AGACAGGGAAGTGGCCAAAACCGTCTTTTGGCTGGTCTTGTCTTTGCC	1236
300	lnaTgArgLyuAlaLysThrValPheCysLeuValIlePheAla	316
1237	CTCTGCTGGCTCCCTTCACCTCAGCAGGATTCTGAAGCTCACTCTTTA	1286
317	LeuCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTy	333
1287	TAATCAGAAATGATCCCAATAGATGTGAACCTTTGAGCTTTCTGTTGGTAT	1336
333	rAspGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuLeuM	350
1337	TGGACTATATTGGTATCAACATGGCTTCACCTGAATTCCTGATTAAACCA	1386
350	eTAspTyrlleGlyIleAsnLeuAlaTrpMetAsnSerCysIleAsnPro	366
1387	ATTGCTCTGTATTTGGTCAGCAAAAGATTCAAAAACCTGCTTTAACTCATG	1436
367	IleAlaLeuTyTrpPheValSerLysLysPheLysAsnCysPheGlnSerCy	383
1437	CTTATGCTGCTGGTGC...CAGTCATTTGAAGAAACAGTCCTTGGAGG	1483
383	sLeuCysCysCysCysTyTrpGlnSerLysSerLeuMetThrSerValProM	400
1484	AAAGACGACTGCTGCTTAAAGTTCAAAGCTCAATGAT...CACGGATATGAC	1530
400	eTAsnGlyTrpSerIleGlnTrpLysAsnHisGluGlnAsnHisAsn	416
1531	AACTTCCCTTCAGTAATAATACAGC	1557
417	ThrGluArgSerSerHisLysAspSer	425
seq_name: SwissProt_40:ETLR_PIG		
seq_documentation_block:		
ID	ETLR_PIG	STANDARD;
AC	Q29010;	PRT;
DT	15-JUL-1998 (Rel. 36, Created)	
DT	15-JUL-1998 (Rel. 36, Last sequence update)	
DT	15-JUL-1998 (Rel. 36, Last annotation update)	
DE	Endothelin-1 receptor precursor (ET-A).	
GN	EDNR.	
OS	Sus scrofa (Pig).	
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.	
OX	NCBI_TaxID=9823;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RC	TISSUE=Lung;	
RX	MEDLINE=96031034; PubMed=8564193;	
RA	Nishimura J., Aoki H., Chen X., Shikasho T., Kobayashi S.,	
RA	Kanade H.;	
RT	"Evidence for the presence of endothelin ETA receptors in endothelial	
RL	cells in situ on the aortic side of porcine aortic valve.";	
CC	-1- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY	
CC	ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-	
CC	CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING	
CC	AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3.	
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein.	
CC	-1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.	
CC	This SWISS-PROT entry is copyright IT is produced through a collaboration	


```

|||||
169 valGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTyrAr 185
840 AGCTGTTGCTCTTGGAGTAAATAAGAGGATGGGCTTCCAAATGGA 889
185 gAlaValAlaSerTrpSerArgValGlnGlyIleGlyLeuProLeuValT 202
890 CAGCAGTAGAATGTTTGGTGGTGGTGGTCTCTGTGGTCTGTGCTGTC 939
202 hrAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAlaIle 218
940 CCTGAAGCCATAGTGTGTATATATACGATGAGTACGACAAAGAGTTA 989
219 ProGluAlaIleGlyPheValMetValProPheGluTyrLysGlyGluI 235
990 TCTCGCAATGCTGCTTTCATCCGCTTCAGAAAGACAGCTTTCATGCAGT 1039
235 uHisLysThrCysMetLeuAsnAlaThrSerLys.....PheMetGluP 250
1040 TTTACAAGACAGCAAAAGATGGTGGTGGTTCAGTTTCTATTTCTGCTG 1089
250 heTyrGlnAspValLysAspTrpIlePheGlyPheTyrPheCysMet 266
1090 CCATTGCCCATCTGCTATTTTATACACTAATGACCTGTGAATGTT 1139
267 ProLeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLe 283
1140 GAGAAAGAAAGTGGC...ATGCAGATGCTTTAAATGATCACCCTAAGC 1186
283 uAsnArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysG 300
1187 AGACAGCGGAAGTGGCCAAACCGCTTTTGGCTGGTCTTGTCTTTGCC 1236
300 lnArgArgGluValAlaLysThrValPheCysLeuValValIlePheAla 316
1237 CTCTGCTGGCTTCCCTTCACCTCAGCAGGATTCCTGAAGCTCCTCTTA 1286
317 LeuCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValT 333
1287 TAATCAGATGATCCCAATAGATGTAACCTTTTTCAGCTTCTGTGGTAT 1336
333 rAspGluMetAspLysAsnArgCysGluLeuLeuSerPheLeuLeuLeu 350
1337 TGGACTATATGGTATCAACATGCTTCTACTGAATTCCTGCATTAACCCA 1386
350 etAspTyrIleGlyIleAsnLeuAlaThrMetAsnSerCysIleAsnPro 366
1387 ATTGCTCTGATTTGGTGACAAAGATTCAAAACCTGCTTAACTGATCATG 1436
367 IleAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCy 383
1437 CTTATGCTGCTGTGTC...CAGTCATTTTGAAGAAACACAGCTCTGGAGG 1483
383 sLeuCysCysCysCysTyrGlnSerLysSerLeuMetThrSerValPro 400
1484 AAAAGCAGTCGCTCTTAAGCTCAAGCTAATGAT...CACGGATATGAC 1530
400 etAsnGlyThrSerIleGlnTrpLysAsnHisGluGlnAsnAsnHisAsn 416
1531 AACTTCGGTTCAGTAAATAATACAGC 1557
417 ThrGluArgSerSerHisLysAspSer 425
seq_name: SwissProt_40:ET3R_XENLA

```

```

seq_documentation_block:
ID ET3R_XENLA STANDARD; PRT; 444 AA.
AC P32940;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-OCT-1993 (Rel. 27, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Endothelin-3 receptor precursor (ET-C).
OS Xenopus laevis (African clawed frog).

```

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Dermal melanophores;
RX MEDLINE=93366838; PubMed=8360195;
RA Karne S., Jayawickreme C.K., Lerner M.R.;
RT "Cloning and characterization of an endothelin-3 specific receptor
(ETC receptor) from Xenopus laevis dermal melanophores.";
RL J. Biol. Chem. 268:19126-19133(1993).
CC -1- FUNCTION: RECEPTOR FOR ENDOTHELIN-3. MEDIATES ITS ACTION BY
ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CALCIUM SECOND MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC
EMBL: L20299; AAA49704.1;
PIR: A48538; A48538.
GCRDB: GCR_0764;
DR InterPro: IPR000276; GPCR_Rhodpsn.
DR Pfam: PF00001; 7tm_1; 2.
DR PRINTS; PR00237; GPCRHHODOPS.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; FALSE_NEG.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
FT CHAIN 1..18
FT DOMAIN 19..444 ENDOTHELIN-3 RECEPTOR.
FT TRANSMEM 89..113 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 114..124 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 125..152 2 (POTENTIAL).
FT DOMAIN 153..161 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 162..180 3 (POTENTIAL).
FT DOMAIN 181..201 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 202..226 4 (POTENTIAL).
FT DOMAIN 227..254 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 255..279 5 (POTENTIAL).
FT DOMAIN 280..307 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 308..334 6 (POTENTIAL).
FT DOMAIN 335..351 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 360..386 7 (POTENTIAL).
FT DOMAIN 387..444 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 60..60 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 444 AA; 50172 MW; 80AB7EC36720D783 CRC64;

```

```

alignment_scores:
Quality: 1073.00 Length: 432
Ratio: 3.417 Gaps: 11
Percent similarity: 72.685 Percent identity: 52.083
alignment_block:
US-09-931-157-2 x ET3R_XENLA
Align seg 1/1 to: ET3R_XENLA from: 1 to: 444

```

```

385 ACTAAGACCTTATGCGCAAGGTTCCAAAGCCAGCTGCGCGGTCGTT 434
|||||
25 ThrGlnGlnAsnPheProAspIleSerAsnProSerGlnGluLeuAsnG 41
435 GGCACCTGCG..... 444
|||||
41 nGluProAlaHisArgIleValGlnLeuAspSerIleGlnAsnGlyA 58

```

```
445 .....GAGCTGCCCTAAAGGACACAGGACGGCAGGATCTCCGCCACGCCACC 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
58 laLeuAsnMetSerThrGlyValAsnValLeuAsnMetSerProPro..... 72
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
490 ATCTCCCTCCCTCCCTGCGCAAGGCCATCCAGATCAGAGAGACTTTTCAA 539
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
73 ...ProProSerProCysLeuSerArgAlaLysIleArgHisAlaPheLeu 88
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
540 ATACATCAACACGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 589
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
88 sTyValThrThrIleLeuSerCysValIlePheLeuValGlyIleValG 105
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
590 GGAATCCACACTTCTTGAGAAATATATACAAAGACACAGTGCATGGAAAC 639
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
105 lYasSerThrLeuLeuArgIleIleTyLysAsnLysCysMetArgAsn 121
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
640 GGTCCCAATATCTGATCGCCAGCTGGCTCTGGGAGACCTGCTGCACAT 689
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
122 GlyProAsnValLeuIleAlaSerLeuAlaLeuGlyAspLeuPheTyri 138
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
690 CTTATTGACATCCCTATCAATGTCTACAAGCTGCTGGCAGAGGACTGCC 739
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
138 leuIleAlaIleProIleIleSerIleSerPheTrpLeuSerThrGly 154
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
740 CATTTGGAGCTGAGATGTGTAGCTGGTGGCTTTTCATACAGAAAGCCCTCC 789
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
155 HisSerGluTyI...IleTyGlnLeuValHisLeuTyArg..... 167
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
790 GTGGGAATCACTGCTGCTGAGCTATGCTGCTGCTGCTGCTGCTGCTGCT 839
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
168 AlaArgValTySerLeuSerLeuCysAlaLeuSerIleAspArgTyArg 184
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
840 AGCTGTTCTCTCTGAGTAGAATTAAGGAATTTGGGTTTCCAAATGGA 889
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
184 gAlaValAlaSerTrpAsnArgIleArgSerIleGlyIleProValArgL 201
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
890 CAGCAGTAGAATTTTGTATTTGGTGGTGGTGGTGGTGGTGGTGGTGGT 939
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
201 ysAlaIleGluLeuThrLeuIleTrpAlaValAlaIleValAlaVal 217
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
940 CCTGAAGCCATAGTTTGTATATATACGATGAGTACAAAGGAAGTTA 989
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
218 ProGluAlaIleAlaPheAsnLeuValGluLeuAspPheArgGlyGln 234
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
990 TCTGCAATCTGCTGCTTCCCTTCCCTTCCAGAGACAGCTTTTCATGCAGT 1039
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
234 rIleLeuValCysMetLeuProMetGluGlnThrSerAspPheMetArgP 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1040 TTTCAAGACAGCAAAAGATTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1089
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
251 heTyGlnGluValLysValTrpTrpLeuPheGlyPheTyPheCysLeu 267
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1090 CCATTGGCCATCAGTCGATTTTATACACTAATGACCTGTAATGTTT 1139
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
268 ProLeuAlaCysThrGlyValPheTyThrLeuMetSerCysGluMetLe 284
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1140 GAGAAGAAAAGTGCATGCGATGCTTTAAATGATCACCCTAAAGCAGA 1189
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
284 uSerIleLysAsnGlyMetArgIleAlaLeuAsnAspHisMetLysGlnA 301
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1190 GACGGAGATGGCCAAACCGTCTTTTGGCTGGTGGTGGTGGTGGTGGTGG 1239
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
301 rArgGluValAlaLysThrValPheCysLeuValValIlePheAlaLeu 317
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1240 TCCTGGCTCCCTTCACCTCAGCAGGATT.....CTGAACCTCAC 1280
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
318 CysTrpLeuProLeuHisValSerSerIlePheValArgLeuSerAlaThr 334
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1281 T.....CTTTAATACAGAAATGATCCCAATAGATGTG 1312
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
334 rValLysArgAlaCysIleLeuLysAsnLysArgSerCysIleMetAlaG 351
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
```

```
1313 AACCTT.....TTGAGCTTTCTGTTGGTATTGGACTATATT 1347
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
351 luIleGlnThrGlyValAsnTyGlnLeuLeuMetValMetAsnTyThr 367
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1348 GGTATCAACATGGCTTCACTGAATTCCTGCAATTAACCAATTCGCTCTGTA 1397
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
368 GlyIleAsnMetAlaSerLeuAsnSerCysIleGlyProValAlaLeuTy 384
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1398 TTTGGTGAGCAAAAGATTCAAAACTGCTTTAAGTCATGCTTATGCTGCT 1447
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
384 rPheValSerArgLysPheLysAsnCysPheGlnSerCysLeuCysCyst 401
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1448 GGTGCCAGTCATTTCAAGAAAAACAGTCCTTG.....GAG 1482
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
401 rPCys.....HisArgProThrLeuThrIleThrProMetAsp 413
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1483 GAAAGCAGTCGCTGCTTAAAGTTCAAGCTAATGATCAGGATATGACAA 1532
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
414 GluLysGlySerGlyLysTrpLysAlaAsnGlyHisAspLeuAspLeu 430
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1533 CTTCGCTTCC.....AGTAATAATACAGCTCATCT 1563
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
430 uAspArgSerSerArgLeuSerAsnLysTySerSerSer 444
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
seq_name: SwissProt_40:ETIR_MOUSE

seq_documentation_block:
ID ETIR_MOUSE STANDARD; PRT; 270 AA.
AC Q61614; O54993;
DT 01-NOV-1997 (Rel. 35, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Endothelin-1 receptor (ET-A) (ET-AR) (Fragment).
GN EDNRA OR GPCR10.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Lung;
RA Schoenfeld J.R., Lowe D.G.;
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 100-216 FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=94116980; PubMed=8288218;
RA Wilkie T.M., Chen Y., Gilbert D.J., Moore K.J., Yu L., Simon M.I.,
RA Copeland N.G., Jenkins N.A.;
RT "Identification, chromosomal location, and genome organization of
RL mammalian G-protein-coupled receptors.";
RL Genomics 18:175-184(1993).
CC -!- FUNCTION: RECEPTOR FOR ENDOTHELIN-1. MEDIATES ITS ACTION BY
CC ASSOCIATION WITH G PROTEINS THAT ACTIVATE A PHOSPHATIDYLINOSITOL-
CC CALCIUM SECOND MESSENGER SYSTEM. THE RANK ORDER OF BINDING
CC AFFINITIES FOR ET-A IS: ET1 > ET2 >> ET3 (BY SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
-----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
CC EMBL; AF039892; AAB96671.1; -.
DR EMBL; L20340; AAA16845.1; -.
DR GCRDB; GCR_0925; -.
DR MGD; MGI:105923; Ednra.
DR MGD; MGI:99570; Gpcr10.
DR InterPro; IPR000276; GPCR_Rhodpsn.
```


DR PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein.
FT NON_TER 1 1
FT DOMAIN <1 8 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 9 30 3 (POTENTIAL).
FT DOMAIN 31 51 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 52 76 4 (POTENTIAL).
FT DOMAIN 77 >99 EXTRACELLULAR (POTENTIAL).
FT DISULFID 7 88 BY SIMILARITY.
FT NON_TER 99 99
SQ SEQUENCE 99 AA; 10943 MW; AA135608FE26BFCA CRC64;

alignment_scores:
Quality: 504.00 Length: 99
Ratio: 5.091 Gaps: 0
Percent Similarity: 100.000 Percent Identity: 97.980
alignment_block:
US-09-931-157-2 x ETBR_MACFA ..
Align seg 1/1 to: ETBR_MACFA from: 1 to: 99
739 CCATTGGAGCTGAGATGCTAAGCTGGTCCCTTTCATACAGAAAGCCTC 788
|||||
1 ProPheGlyAlaGluMetCysLeuValProPheIleGlnLysAlaSe 17
789 CGTGGGAATCACTGCTGAGTCTATGCTCTGAGTATTGACAGATATC 838
|||||
17 rValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrA 34
839 GAGCTGTTGCTCTTTGGAGTAGAATAAGGAATTGGGCTTCCAAAATGG 888
|||||
34 rGAlaValAlaSerTrpSerArgIleLysGlyIleGlyIleProLysTrp 50
889 ACAGCAGTAGAAATGTTTGGTGGTGGTCTCTCTGCTGCTTCTGCCTGT 938
|||||
51 ThrAlaValGluIleValLeuIleTrpValValSerValValLeuAlaVa 67
939 CCTGAAGCCATAGGTTTTGTATATAATTACGATGGACTACAAAGGAAGTT 988
|||||
67 lProGluAlaIleGlyPheAspMetIleThrMetAspTyrLysGlySerT 84
989 ATCTCGGAATCTGCTTTCATCCCGTTTCAGACAGACAGCTTTCATG 1035
|||||
84 yrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMet 99

OM of: US-09-931-157-2 to: PIR_71:* out_format : pfs

Date: Oct 1, 2002 1:11 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

```
-MODEL=frame+n2p.model -DEV=xlp
-Q/Cgn2_1/USPTO.spool/US09931157/runat_30092002.110421.11282/app_query.fasta_1.8657
-DB=PIR_71 -QFMT=fastan -SUFFIX=xpr -CAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -CAPOP=4.500
-GAPEXT=0.050 -XCAPOP=10.000 -XGAPEXT=0.500 -FGAPEXT=6.000
-DEPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELXT=7.000 -START=1 -MATRIX=blosom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEAPSIZ=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09931157 -CGN1.1.335
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NOLXPY -WAIT -THREADS=1
```

Search information block:

Query: US-09-931-157-2

Query length: 4301

Database: PIR_71:

Database sequences: 283138

Database length: 96089334

Search time (sec): 273.170000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	Documentation
pir1:JQ1042	+ 2340.00	3603.26	2.9e-193	442	! endothelin receptor B precursor
pir2:I57950	+ 2070.00	3185.20	5.7e-170	442	! nonselective-type endothelin re
pir2:A41591	+ 2046.50	3148.83	6.0e-168	441	! endothelin receptor B precursor
pir2:S13425	+ 2044.50	3145.74	9.0e-168	441	! endothelin receptor ETB - rat
pir2:A40440	+ 1245.50	1908.91	7.3e-99	426	! endothelin 1 and 2 receptor pre
pir2:A54126	+ 1244.50	1907.60	8.8e-99	415	! endothelin receptor-AX - Africa
pir2:A44158	+ 1240.00	1900.37	2.2e-98	427	! endothelin receptor type A pred
pir2:S13424	+ 1238.00	1897.28	3.2e-98	427	! endothelin receptor - bovine
pir2:A48538	+ 1073.00	1641.43	5.5e-84	444	! endothelin-3 receptor precursor
pir2:A39003	+ 498.50	753.24	1.9e-34	384	! bombesin/gastrin-releasing pep
pir2:I57682	+ 496.50	750.14	2.8e-34	384	! bombesin/ GRP receptor - rat
pir2:B41007	+ 491.00	741.48	8.4e-34	390	! bombesin receptor, neuromedin B
pir2:JH0374	+ 489.50	739.16	1.1e-33	390	! bombesin receptor, neuromedin B
pir2:A41007	+ 466.00	702.92	1.2e-31	384	! gastrin-releasing peptide recep
pir2:A46632	+ 443.00	666.95	1.2e-29	399	! bombesin-like peptide receptor
pir2:I50102	+ 441.50	665.18	1.6e-29	376	! Phel3 bombesin receptor - Bombi
pir2:S29480	+ 441.00	663.85	1.7e-29	399	! bombesin receptor - guinea pig
pir2:B57508	+ 402.50	615.58	2.9e-26	117	! endothelin-1 receptor homolog -
pir2:JC5501	+ 385.50	573.95	1.1e-24	613	! endothelin receptor type-B-like
pir2:JC7209	+ 348.50	520.75	1.6e-21	394	! galanin receptor - fruit fly (D
pir2:JN0605	+ 331.50	494.56	4.8e-20	388	! somatostatin receptor 4 - human
pir2:JC2338	+ 329.50	491.66	7.1e-20	380	! kappa opioid receptor 1 - human
pir2:A41795	+ 326.50	484.40	1.4e-19	504	! tachykinin receptor NKD - fruit
pir2:I39187	+ 324.00	482.88	2.1e-19	391	! somatostatin receptor 1 - human
pir2:C41795	+ 323.50	482.35	2.3e-19	381	! neurotensin Y-peptide yv recep
pir2:JC2434	+ 322.00	479.78	3.2e-19	391	! somatostatin receptor 1 - mouse
pir2:A48227	+ 321.50	479.27	3.5e-19	380	! kappa opioid receptor - mouse
pir2:JC7677	+ 321.50	479.27	3.5e-19	380	! kappa opioid receptor - guinea
pir2:S36143	+ 320.50	476.73	4.3e-19	423	! allatostatin receptor - America
pir2:A47249	+ 318.50	474.63	6.3e-19	380	! kappa opioid receptor - rat
pir2:JC7319	+ 316.50	471.44	9.4e-19	394	! brain-specific somatostatin rec
pir2:A39297	+ 311.00	463.59	2.8e-18	357	! probable allatostatin receptor
pir2:S43087	+ 309.50	460.43	3.8e-18	391	! somatostatin receptor - rat
pir2:S65693	+ 308.00	458.62	5.1e-18	370	! orphan opioid receptor ORL1 -
pir2:J01231	+ 308.00	458.08	5.1e-18	352	! opioid receptor mu variant MOR1
pir2:I59336	+ 307.50	458.23	5.5e-18	355	! interleukin-8 receptor - rabbit
pir2:I56553	+ 307.00	457.61	6.1e-18	349	! galanin receptor 1 - human
pir2:I56517	+ 307.00	456.35	6.3e-18	398	! mu opiate receptor - human
pir2:I56517	+ 304.50	452.52	1.0e-17	400	! mu opioid receptor - rat

```
pir2:I56520 + 304.00 452.50 1.1e-17 367 ! G protein-coupled recept
pir2:JC2083 + 303.50 451.67 1.2e-17 369 ! somatostatin receptor 2
pir2:JC4162 + 303.50 451.57 1.2e-17 373 ! p2y receptor - bovine
pir2:JC2421 + 303.00 450.95 1.4e-17 367 ! opioid receptor homolog,
pir1:I52315 + 303.00 450.88 1.4e-17 370 ! G protein-coupled recept

seq_name: pir1:JQ1042
seq_documentation_block:
endothelin receptor B precursor - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 21-Jul-2000
C:Accession: A46609; JQ1042; JN0072; JQ1082; B45481
R:Arai, H.; Nakao, K.; Takaya, K.; Hosoda, K.; Ogawa, Y.; Nakanishi, S.; Imura, H
J. Biol. Chem. 268, 3463-3470, 1993
A:Title: The human endothelin-B receptor gene. Structural organization and chromo
A:Reference number: A46609; MUID:93155196
A:Accession: A46609
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-442 <ARA>
A:Cross-references: GB:D13168; NID:q285924; PIDN:BA02445.1; PID:q285926
R:Ogawa, Y.; Nakao, K.; Arai, H.; Nakagawa, O.; Hosoda, K.; Suga, S.; Nakanishi,
Biochem. Biophys. Res. Commun. 178, 248-255, 1991
A:Title: Molecular cloning of a non-isopeptide-selective human endothelin recepto
A:Reference number: JQ1042; MUID:91298956
A:Accession: JQ1042
A:Molecule type: mRNA
A:Residues: 1-442 <OGA>
A:Cross-references: GB:D90402; NID:q219651; PIDN:BA014398.1; PID:q219652
R:Nakamura, M.; Takayanagi, R.; Sakai, Y.; Sakamoto, S.; Hagiwara, H.; Mizuno, T.
Biochem. Biophys. Res. Commun. 177, 34-39, 1991
A:Title: Cloning and sequence analysis of a cDNA encoding human non-selective typ
A:Reference number: JN0072; MUID:91254298
A:Accession: JN0072
A:Molecule type: mRNA
A:Residues: 1-442 <NAK>
A:Cross-references: GB:M74921; NID:g182275; PIDN:AAA58465.1; PID:g182276
R:Experimental source: liver
R:Sakamoto, A.; Yanagisawa, M.; Sakurai, T.; Takawa, Y.; Yanagisawa, H.; Masaki,
Biochem. Biophys. Res. Commun. 178, 656-663, 1991
A:Title: Cloning and functional expression of human cDNA for the ETB endothelin r
A:Reference number: JQ1082; MUID:91315496
A:Accession: JQ1082
A:Molecule type: mRNA
A:Residues: 1-9, 'p', 11-442 <SAK>
A:Cross-references: GB:S44866; NID:q233233; PIDN:AA019411.1; PID:q233234
R:Experimental source: jejunum
R:Elshourbagy, N.A.; Korman, D.R.; Wu, H.L.; Sylvestre, D.R.; Lee, J.A.; Nuthalag
J. Biol. Chem. 268, 3873-3879, 1993
A:Title: Molecular characterization and regulation of the human endothelin recept
A:Reference number: A45481; MUID:93179382
A:Accession: B45481
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-442 <ELS>
A:Cross-references: GB:L06623; NID:g181958; PIDN:AAA52342.1; PID:g181959
R:Experimental source: lung
A:Note: sequence extracted from NCBI backbone (NCBIN:125754, NCBI:125755)
C:Comment: This endothelin receptor is non-isopeptide selective.
C:Genetics:
A:Gene: GDB:EDNRB; ETB
A:Cross-references: GDB:129075; OMIM:131244
A:Map position: 13q22-13q22
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembra
F:1-26/Domain: signal sequence #status predicted <SIG>
F:102-127/Domain: transmembrane #status predicted <TM1>
F:138-163/Domain: transmembrane #status predicted <TM2>
F:176-197/Domain: transmembrane #status predicted <TM3>
F:219-243/Domain: transmembrane #status predicted <TM4>
```

F:272-296/Domain: transmembrane #status predicted <TM5>
 F:325-350/Domain: transmembrane #status predicted <TM6>
 F:363-389/Domain: transmembrane #status predicted <TM7>
 F:59/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:305,413,419,435,440/Binding site: phosphate (Ser) (covalent) #status predicted

alignment_scores:

Quality: 2340.00 Length: 442
 Ratio: 5.294 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000

alignment_block:

US-09-931-157-2 x JQ1042

Align seg 1/1 to: JQ1042 from: 1 to: 442

```

238 ATGCAGCGCCCTCAAGTCTGTGGGACGCCCTGGTGGCGGTCTTCT 287
|||||
1 MetGlnProProProSerLeuCysGlyArgAlaLeuValAlaLeuVal 17
288 TGCCTGGCGCTGTCGGGATCTGGGAGAGGAGAGAGGCTTCCCGCTG 337
|||||
17 uAlaCysGlyLeuSerArgIleTrpGlyGluGluArgGlyPheProPro 34
338 ACAGGGCCACTCCGCTTTTGCACCGCAGAGATAATGACGCCACCCACT 387
|||||
34 sPArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProProThr 50
388 AAGACCTTATGGCCCAAGGTTTCCAAACGCCAGTCTGGCGGCTGTGGC 437
|||||
51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeuAl 67
438 ACCTGGCGAGTGCTTAAAGGACAGACGCGCAGGATCTCGGCACCGCA 487
|||||
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArgT 84
488 CCATCTCCCTCCCGCTGCAAGGACCCATCGAGATCAAGGAGACTTTC 537
|||||
84 hrlleSerProProProCysGlnGlyProIleGluIleLysGluThrPhe 100
538 AAATPACATCAACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 587
|||||
101 LysThrIleAsnThrValValSerCysLeuValPheValLeuGlyIleI 117
588 CGGGAATCCACACTCTGAGAAATATCTACAAGACAAAGTCATGCCAA 637
|||||
117 eGlyAsnSerThrLeuLeuArgIleIleIleIleIleIleIleIleIle 134
638 ACGGTCCCAATATCTTGATCGCCAGCTTGGCTCTGGGAGACCTGCTGC 687
|||||
134 sNGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuHis 150
688 ATCTCATATTGACATCCCTATCAATGTCTACAAGCTGCTGGCAGAGACT 737
|||||
151 IleValIleAspIleProIleAsnValTrpLysLeuLeuAlaGluAspTr 167
738 GCCATTTGGAGCTGAGATGTAGTGGTGGCTTTCATACAGAAAGCTT 787
|||||
167 pPropheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGTGGGAATCACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
|||||
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
838 CGAGCTGTGCTTCTGGAGTAGAATTAAGGAATTTGGGGTTCCAAAATG 887
|||||
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
888 GACAGCAGTAGAAATTTGTTGATTTGGTGGTGGTCTCTGCTGCTGCTG 937
|||||
217 pThrAlaValGluIleValLeuIleTrpValValSerValValLeuAlav 234

```

```

938 TCCCTGAAGCCATAGTGTGATATATATACGATGACGACTACAAAGAGT 987
|||||
234 alProGluAlaIleGlyPheAspIleIleThrMetAspTrpLysGlySer 250
988 TATCTGCGAATCTGCTGCTCATCCCGTTTCAGAAAGACAGCTTTCATG 1037
|||||
251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetG 267
1038 GTTTTACAGACAGCAAAAGATTGTGGCTGTTCAGTTCTTATTTCTGCT 1087
|||||
267 nPheTrpLysThrAlaLysAspTrpIlePheSerPheTrpPheCysL 284
1088 TGCATTGGCCATCAGTCTGATTTTATATACACTAATGACCTGTAAGT 1137
|||||
284 euProLeuAlaIleThrAlaPheTrpThrLeuMetThrCysGluMet 300
1138 TTGAGAAAGAAAGTGGCAGTGTGCTTTAAATGATGATCACCATAAGCA 1187
|||||
301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
1188 GAGCGGGAAGTGGCCAAACCGTCTTTGCTGCTGCTGCTGCTGCTGCT 1237
|||||
317 nArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAl 334
1238 TCTGCTGGCTTCCCTTCACTCAGCAGGATTTGAAAGCTCACTCTTTAT 1287
|||||
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTr 350
1288 AATCAGAATGATCCCAATAGATGTGAACCTTTGAGCTTTCTGTTGTATT 1337
|||||
351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuVal 367
1338 GGACTATATTGCTATCAACATGGCTTCACTGATTCCTGCTGATTAACCC 1387
|||||
367 uAspTrpIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 384
1388 TTGCTCTGTATTGTTGGTGAAGAAAGATTCAAAACCTGCTTTAAGTATGC 1437
|||||
384 leAlaLeuTrpLeuValSerLysArgPheLysAsnCysPheLysSerCys 400
1438 TTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1487
|||||
401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGlu 417
1488 CCAGTCTGCTGCTTAAAGTTCARAGCTAATGATCAGGATATGACACTTC 1537
|||||
417 sGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnPhe 434
1538 GTTCAGTAAATAATACAGCTCATCT 1563
|||||
434 rgSerSerAsnLysTrpSerSerSer 442

```

seq_name: pir2:157950

seq_documentation_block:

nonselective-type endothelin receptor - rat

C:Species: Rattus sp. (rat)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999

C:Accession: I57950

R:Cheng, H.F.; Su, Y.M.; Yeh, J.R.; Chang, K.J.

Mol. Pharmacol. 44, 533-538, 1993

A:Title: Alternative transcript of the nonselective-type endothelin receptor from

A:Reference number: I57950; MUID:93382424

A:Accession: I57950

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-442 <RES>

A:Cross-references: GB:S65355; NID:g410692; PIDN:AA828172.1; PID:g410693

C:Superfamily: endothelin receptor B

alignment_scores:

Quality: 2070.00

Length: 443

Eur. J. Biochem. 234, 251-257, 1995

A:Title: Isolation of the endothelin B receptor from bovine lung. Structure, signal sequence and function of the receptor

A:Reference number: S63513; MUID:96096746

A:Accession: S63513

A:Molecule type: protein

C:Residues: 1, 'X', 3-7, 'X', 9-10:25-35;36-42:73-78;79-87;284-290:291-296, 'X', 298 <HIC>

C:Superfamily: endothelin receptor B

C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein

F:1-26/Domain: signal sequence #status experimental <SIG>

F:27-441/Product: endothelin receptor B #status experimental <SIG>

F:101-126/Domain: transmembrane #status predicted <TM1>

F:137-162/Domain: transmembrane #status predicted <TM2>

F:175-196/Domain: transmembrane #status predicted <TM3>

F:218-242/Domain: transmembrane #status predicted <TM4>

F:271-295/Domain: transmembrane #status predicted <TM5>

F:324-349/Domain: transmembrane #status predicted <TM6>

F:362-388/Domain: transmembrane #status predicted <TM7>

F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:

Quality: 2046.50 Length: 443

Ratio: 4.919 Gaps: 2

Percent Similarity: 93.905 Percent Identity: 88.488

alignment_block:

US-09-931-157-2 x A41591 ..

Align seg 1/1 to: A41591 from: 1 to: 441

```

238 ATGACGGCGCTCAAGTCTGTGGGACGGCCCTGTGGTGGTGTCT 287
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
1 MetGlnProLeuProSerLeuCysGlyArgAlaLeuValAlaLeuLe 17
288 TGCTGGCGCTGTCGGGAGTCTGGGAGAGGAGAGGCTTCCCGCTG 337
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
17 uAlaCysGlyValAlaGlyLeGlnAlaGluGluArgGluPhePro 34
338 ACAGGGCCACT...CCGCTTTTGCACACCGCAGAGATAATGACCC 384
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
34 laGlyAlaThrGlnProLeuProGlyThrGlyGluMetMetGluThr 50
385 ACTAAGACCTTATGCGCCAGGTTTCCAGCCAGCTGTGGCGGTGTT 434
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
51 ThrGluThrSerTrpProGlyArgSerAsnAlaSerAspProArgSer 67
435 GGCACCTGCGGAGTGTCTAAAGGAGACAGCGGAGGATCTCCGCCAC 484
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
67 rAlaThrProGlnProArgGlyArgMetAlaGlyLeuProPro 84
485 GCACCACTCCCTCCCGCCGCAAGGACCCATCGAGATCAAGGAGACT 534
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
84 rgThr.....ProProCysaspGlyProIleGluIleLysGluThr 98
535 TTCAATACATACACAGGTGTGCTGCTGTGCTGTGCTGTGCTGTGCT 584
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
99 PheLysTyrIleAsnThrValValSerCysLeuValPheValLeuGly 115
585 CATTGGGAGTCCACACTTCTGAGAAATATCTACAGAAACAGTCATGC 634
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
115 eIleGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMet 132
635 GAAACGGTCCCAATATCTGATCGCCAGCTTGGCTGTGGGAGACTGCTG 684
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
132 rgAsnGlyProAsnIleLeuLeuAlaSerLeuAlaLeuGlyAspLeu 148
685 CACATCGTATGACATCCCTATCAATGCTTACAGCTGTGCGCAGAGGA 734
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
149 HisIleIleIleAspIleProIleAsnThrTyrLysLeuLeuAlaLys 165
735 CTGGCCATTTGGAGCTGAGATGTAGCTGGTGGTGGTGGTGGTGGTGG 784
|||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
165 pTrpProPheGlyValGluMetCysLysLeuValProPheIleGlnLys 182

```

seq_name: pir2:sl3425

seq_documentation_block:

endothelin receptor ETB - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 16-Jul-1999

C:Accession: S13425; A49197

R:Sakurai, T.; Yanagisawa, M.; Takuwa, Y.; Miyazaki, H.; Kimura, S.; Goto, K.; Mas

165 pTrpPropheClyAlaGluMetCysLysLeuValProPheIleGlnLysA 182
785 CCTCCGCGGAATCACGTGCTGAGTCTATCTGCTCAGTATTACACAGA 834
182 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleLeasparg 198

835 TATCGAGCTGTGCTCTCTTGGAGCTAGAATTAAGGAATTTGGGGTTCAAA 884
 199 TTTArGAlaValAlaSerTrpSerArgIleLySGlyIleGlyValProLy 215
 885 ATGGACAGCAGTAGAAAATGTTTTGATTTGGGTGGTCTCTGTGGGTCTGG 934

[illegible]

035 GCAGTTTACAAGACAGCAAAAGATTGGTGCTTCAGTTCTATTTCT 108
249 LysProLeuArgValCysMetLeuAsnProPheGlnHisIleAlaPheMe 266

11085 GCTTGCCATTGGCCATCACTGCATTTTTTATACACTAATGACCTGTGAA 113

282 ysLeuProLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGlu 298

[illegible]

299 MetLeuArgLysLysSerGlyMetValIleTyrAlaLeuAsnAspHisLeuVal 133
105 CAC 133

11185 GCAGAGACGGGAAGTGGCCAGAACCGTCTTTTGGCTGGTCTGCTCTTG 122

315 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 332

11235 CCCTCTGCTGGCTTCCCCCTCAGCAGGATTCTGAAGCTCACCTT 128

332 1aLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 348

11285 TATAATCAGAATGATCCCAATAGATGTGAACTTTTGAGCTTTCTGTTGGT 133

349 TyrAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeuVa 365

1335 ATGGACTATATTGGTATCAACATGGCTTCACTGAATTCCTGCATTAAACC 138

365 iLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 382

1385 CAATTGCTCTGTATTTGGTGAGCAAAAGATTCAAAAACGCTTTAAGTCA 143

382 roilleAlaLeuTyrLeuValSerLysArqPheLysAsnCysPheLysSer 398

1435 TGCTTATGCTGCTGGTCCAGTCATTTGAAGAAAAACAGTCCTTGGAGGA 148

[illegible]

399 cysteucysstipcysgmimimfneegluglulysgmimseileuglugl 411

11485 AAAGCAGTCGGTCTTAAAGTTCAAAGCTAATGATCACGGATATGACAACT 153

415 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyTrAspAsnP 432

1535 TCCGTTCCAGTAATAAATACAGCTCATCT 1563

432 heArgSerSerAsnLysTyrSerSerSer 441

name: pir2:A40440

```
documentation_block:
```

thelin 1 and 2 receptor precursor - rat
species: Rattus norvegicus (Norway rat)

```
te: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change
session: A40440; B49197
```

.....

name: pir2:A40440

documentation block:

ethelin 1 and 2 receptor precursor - rat

Species: *Rattus norvegicus* (Norway rat)

```
te: 22-Jan-1993 #sequence_revision 22-Jan-1993 #text_change
session: A40440; B49197
```


alignment_scores:

Quality: 1244.50 Length: 428
Ratio: 3.760 Gaps: 6
Percent Similarity: 77.336 Percent Identity: 56.542

alignment_block:

US-09-931-157-2 x A54126

Align seg 1/1 to: A54126 from: 1 to: 415

```
265 CGCGCCCTGGTTCGCTGCTTCTTGGCC...TGGCGGCTGTCGGGATCTG 311
|||||
7 ArgPheThrValLeuLeuValLeuAlaGlyIleAlaValSerSerPh 23
312 GGAGAGGAGAGAGGCTCCCGCTGACAGGGCCACTCCGCTTTCCAAA 361
|||||
23 eGlyGluTyrTyrGlnAsnArgThrAspAlaSerThrMetL 40
362 CGCAGAGATATATGACGCCACCCACTAAGACCTTATGGCCCAAGGTTCC 411
|||
40 euAsnArgSerHisThrSerPro..... 47
412 AACCCCACTGTCGCCGCTGCTTGGCACCTGCGGAGGTGCCTAAAGAGA 461
|||
48 .....ValArgLysGlyAs 52
462 CAGGACGGCAGGATCTCCGCCAGCACCATCTCCCTCCCGCTGCCAAG 511
|||
52 narGSerAlaAspLeu.....CysProG 60
512 GACCCATCAGATCAAGGAGACTTTCAAATACATCAACACAGGTGTGTC 561
|||||
60 lulysthrLysIleAsnHisValPheLysTyrIleAsnThrIleLeuSer 76
562 TGCCTGTGCTGCTGCTGGGATCATCGGAACCTCCACACTTCTGAGAT 611
|||
77 CysThrIlePheIleIleGlyMetValGlyAsnAlaThrLeuLeuArgI 93
612 TATCTACAAGACAAGTGTAGTGGCAACGTCCTCAATATCTTGTATGCC 661
|||||
93 eileTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeuIleAla 110
662 GCTTGGCTGCGGAGACCTGCTGCATCATCTGATTCATTCATCAAT 711
|||||
110 erLeuAlaLeuGlyAspLeuIleTyrIleValIleAspIleProIle 126
712 GTCTACAAGCTGCTGCCAGAGGACTGCCCATTT.....GG 746
127 ValPheLeuLeuLeuGlyGlnArgTyrProPheAspGlnSerProValG 143
747 AGCTGAGATGTGTAAGTGTGCTGCTTTCATACAGAAAGCCTCGTGGGA 796
143 yAlaPheLeuCysLysLeuValProPheIleGlnLysAlaSerValGly 160
797 TCAGTGTGCTGACTCTATGCTGCTGATGATTCACAGATATCGAGCTGT 846
160 lethrValLeuAsnLeuCysAlaLeuSerValAspArgTyrArgAlaVal 176
847 GCTTCTTGGAGTGAATTAAGGAANTTGGGTTCCCAAAATGGACACAGT 896
177 ProSerTrpSerArgValGlnGlySerGlyIleProLeuIleThrAla 193
897 AGAAATGTTTTCATTTGGTGGTCTGCTGCTGCTGCTGCTGCTGCTG 946
193 eGluIleIleSerIletrpValLeuSerPheValLeuAlaIleProGlu 210
947 CCATAGTGTGTGATAAATACCATGGACTACAAAGGAAGTATCTCGCA 996
210 laileGlyPheValMetValProPheGluTyrArgGlyGluGlnPheArg 226
997 ATCTGCTTGTCTTCCCGTTCAAGACAGACAGCTTTCATGCAAGTTTACA 1046
227 ThrCysMetPheHisAlaThr.....SerProPheMetMetPheTyr 241
```

```
1047 GACAGCAAAAGATTGGTGGCTGTTCAGTTTCTATTCTTCTGCTGCCATTGG 1096
|||||
241 sasnaAlaLysAspTrpTrpLeuPheGlyLeuTyrPheCysValProLeuA 258
1097 CCATCACTGCATTTTTTTATACACTAATGACCTGTGAAATGTTGAGAAAG 1146
|||
258 laCysThrGlyValPheTyrThrMetMetThrCysGluMetLeuHisGln 274
1147 AAAAGTGGC...ATGCAGATTGCTTTAAATGATCACCCTAAAGCAGAGACG 1193
|||||
275 ArgLysGlySerLeuArgIleAlaLeuSerGluHisLeuLysGlnArgAr 291
1194 GGAAGTGGCCAAACCGTCTTTTTCCTGGTCTGCTTGTCTTGTGCTCTGCT 1243
|||||
291 gGluValAlaLysThrValPheCysLeuValValIlePheAlaLeuCys 308
1244 GCGTTCCTCCCTTCACCTCAGCAGGATCTGAGCTCCTCTTTTATATCAG 1293
|||||
308 rpPheProLeuHisLeuSerArgIleIleLysAsnThrIleTyrAsnGlu 324
1294 AATGATCCCAATAGATGTGAACCTTTGAGCTTCTGTGTGATTGGACTA 1343
|||||
325 LeuAspProSerArgCysGluLeuLeuSerPheLeuLeuValMetAspPh 341
1344 TATTGTTATCAACATGGCTTCACTGAATTCCTGCTTCAACCAATGCTC 1393
|||||
341 eileSerIleAsnLeuAlaAlaLeuAsnSerCysIleAsnProIleAlaL 358
1394 TGTATTGGTGACCAAAAGATTCAAAACTGCTTTTAAAGTCATGCTTATGC 1443
|||||
358 eufyrPheValSerLysLysPheLysAsnCysPheGlnSerCysLeuCys 374
1444 TGCCTGTGCGCAGTCATTTGAAGAAAACAGTCCTTGGAGGAAAAGCAGTC 1493
|||
375 CysCysCysGlnSerLysThrHisIleAsnThrAlaProMetAsnValTh 391
1494 GTCTTAAAGTTCAAAGCTAATGATCAGCATAT 1527
391 rSerIleGlnTrpLysAsnHisAspGlnAsnTyr 402
```

seq_name: pir2.A44158

seq_documentation_block:

endothelin receptor type A precursor - human
N:Alternate names: endothelin 1 receptor; endothelin-A receptor
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 21-Jul-2000
C:Accession: A44158; S16980; J0972; J0607; P1245; A45481; S6289
R:Hosoda, K.; Nakao, K.; Tamura, N.; Arai, H.; Ogawa, Y.; Suga, S.; Nakanishi, S.
J. Biol. Chem. 267, 18797-18804, 1992
A:Title: Organization, structure, chromosomal assignment, and expression of the g
A:Reference number: A44158; MUID:92406798
A:Accession: A44158
A:Molecule type: DNA
A:Residues: 1-427 <HOS>
A:Cross-references: GB:D11145
A:Note: sequence extracted from NCBI backbone (NCBI:113475)
R:Hosoda, K.; Nakao, K.; Arai, H.; Suga, S.; Ogawa, Y.; Mukoyama, M.; Shirakami,
FEBS Lett. 287, 23-26, 1991
A:Title: Cloning and expression of human endothelin-1 receptor cDNA.
A:Reference number: S16980; MUID:91348221
A:Accession: S16980
A:Molecule type: mRNA
A:Residues: 1-427 <HOS>
A:Cross-references: GB:X61950; MID:g288312; PIDN:CAA43953.1; PID:g288313
R:Cyr, C.; Huebner, K.; Druck, T.; Kris, R.
Biochem. Biophys. Res. Commun. 181, 184-190, 1991
A:Title: Cloning and chromosomal localization of a human endothelin eta receptor.
A:Reference number: J0972; MUID:92068188
A:Accession: J0972
A:Molecule type: mRNA
A:Residues: 1-427 <CYR>

```

alignment_scores:
  Quality: 1240.00      Length: 359
  Ratio: 3.987          Gaps: 5
  Percent Similarity: 86.630  Percent Identity: 63.510

alignment_block:
  US-09-931-157-2 x A44158  ..

Align seg 1/1 to: A44158 from: 1 to: 427

505  TCCCAAGGACCCATCGAGATCAAGGAGACTTTCAAATACATCAACGGT 554
|||||
69  CysProGlnGlnThrIvsThrSerAlaPheIvsTvrlIeAsnThrVa 85

```

1337 TGGACTATATTGGTATCAACATGCGTTTCACCTGAATTCCTGCATTAACCCA 1386
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||
350 etASPTyriIeGlyileAsnLeuAlaThrMetAsnSerCysIleAsnPro 366
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

1387 ATTCTCTGTATTGGTGACAAAGATTCAAAACATGCTTTAAGTCATG 1436
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

367 IleAlaLeuTyriPheValSerLysLysPheLysAsnCysPheIcnsIerCy 383
:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||

F:78-97/Domain: transmembrane #status predicted <TR2>

```

seq_name: pir2:A39003

seq_documentation_block:
  bombesin/gastrin-releasing peptide receptor [similarity] - mouse
  C:Species: Mus musculus (house mouse)
  C:date: 14-Feb-1992 #sequence_revision 14-Feb-1992 #text_change
  C:Accession: A39003; A36553
  R:Battay, J.F.; Way, J.M.; Cortay, M.H.; Shapira, H.; Kusano, K.
  Proc. Natl. Acad. Sci. U.S.A. 88, 395-399, 1991
  A:title: Molecular cloning of the bombesin/gastrin-releasing pep
  A:Reference number: A39003; MUID:91110536
  A:Accession: A39003
  A:Molecule type: mRNA
  A:Residues: 1-384 <BAT>
  A:Cross-references: GB:M57922; NID:g996089; PIDN:AAA75650.1; PID
  A:Note: the sequence in GenBank entry MUSGRPBM, release 113.0,
  R:Spindel, E.R.; Giladi, E.; Brehm, P.; Goodman, R.H.; Segerson,
  Mol. Endocrinol. 4, 1956-1963, 1990
  A:title: Cloning and functional characterization of a complement
  A:Reference number: A36553; MUID:91187004
  A:Accession: A36553
  A>Status: not compared with conceptual translation
  A:Molecule type: mRNA
  A:Residues: 1-308, 'R', 310-384 <SPI>
  A:Cross-references: GB:M61000; GB:M35328; NID:g193650; PIDN:AAA3
  C:Superfamily: endothelin receptor B
  C:Keywords: G protein-coupled receptor; glycoprotein; neurotrans
  F:40-63/Domain: transmembrane #status predicted <TRI>
  F:78-97/Domain: transmembrane #status predicted <TRI>

```

F:116-137/Domain: transmembrane #status predicted <TR3>
 F:154-174/Domain: transmembrane #status predicted <TR4>
 F:210-235/Domain: transmembrane #status predicted <TR5>
 F:267-286/Domain: transmembrane #status predicted <TR6>
 F:299-326/Domain: transmembrane #status predicted <TR7>
 F:20,24,59,191,347/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:
 Quality: 498.50 Length: 362
 Ratio: 2.149 Gaps: 8
 Percent Similarity: 64.088 Percent Identity: 33.702

alignment_block:
 US-09-931-157-2 x A39003

Align seg 1/1 to: A39003 from: 1 to: 384

```

484 CGCACATCTCCCTCCCGCTGGCAAGACCCATCGAGATCAAGAGAC 533
      ::::::::::::::::::::
25  GlnSerLeuSerProProlsMetAspAsn.....TrpPheHisProGl 39
      ::::::::::::::::::::
534 TTTCAATAATACATACACACGGTTGTGCTGCTGCTGCTGCTGCTGCTG 583
      ::::::::::::::::::::
39  yPheIleTyValIleProAlaValTyrglyLeuIleValIleGlyL 56
      ::::::::::::::::::::
584 TCATCGGGAACCTCCACATCTCGAGAAATATCTACAAGAACAGTCATG 633
      ::::::::::::::::::::
56  euIleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 72
      ::::::::::::::::::::
634 CGAAGCGTCCCAATATCTGTATCGCCAGCTGGCTCGCTGGGAGACCTGT 683
      ::::::::::::::::::::
73  ArgAsnValProAsnLeuPheIleSerLeuAlaLeuGlyAspLeuLe 89
      ::::::::::::::::::::
684 GCACATCGCTGATGACATCCCTATCAATCTCTACAGCTGCTGGCAGAG 733
      ::::::::::::::::::::
89  LeuLeuValThrCysAlaProValAspAlaSerLysTyrrLeuAlaAsp 106
      ::::::::::::::::::::
734 ACTGGCCATTGGAGCTGAGATGTGTAGCTGGTGCCTTCATACAGAAA 783
      ::::::::::::::::::::
106 rgTrpLeuPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLe 122
      ::::::::::::::::::::
784 GCTCGCTGGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 833
      ::::::::::::::::::::
123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAsp 139
      ::::::::::::::::::::
834 ATATCCAGCTGTGCTCTCTGGAGTAGAATTAAGGAATGGGTGCCAA 883
      ::::::::::::::::::::
139 gTyrrLysAlaIleValArgProMetAspIleGlnAlaSerHisAlaLeu 156
      ::::::::::::::::::::
884 AATGGACAGCAGTAGAAAATTTGATTTGGTGTGCTCTGCTGCTGCTG 933
      ::::::::::::::::::::
156 etLysIleCysLeuLysAlaAlaLeuIleTrpIleValSerMetLeuLeu 172
      ::::::::::::::::::::
934 GCTGCTCCCTGAGCCATAGTTTGTATATAAATACAGTACTACAAAGG 983
      ::::::::::::::::::::
173 AlaIleProGluAlaValPheSerAspLeuHisProPheHisValLysAs 189
      ::::::::::::::::::::
984 AGTTTATCTCGCAATC.....TGCTTGTCTCATCCCTGCTGAGAAGAC 1027
      ::::::::::::::::::::
189 pThrAsnGlnThrPheIleSerCysAlaProTyrrProHisSerAsnGlu 206
      ::::::::::::::::::::
1028 CTTTTCATGTCAGTTTACAAAGACAGCAAAAGATTGGTGGCTGTTCATG 1077
      ::::::::::::::::::::
206 euHisProLysIleHisSerMetAlaSer.....PheLeuVal 218
      ::::::::::::::::::::
1078 TATTTCCTGCTGTCATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1127
      ::::::::::::::::::::
219 PheTyrrValIleProLeuAlaIleIleSerValTyrrTyrrPheIleAl 235
      ::::::::::::::::::::
1128 CTCTGAAATCTTCAGAAAGAAAGTGGCATGTCAGATTGCTTTAAATGATC 1177
      ::::::::::::::::::::
235 aArgAsnLeuIleGlnSerAlaTyrrAsnLeuProValGluGlyAsnIle 252
      ::::::::::::::::::::

```

```

1178 ACTTAAAGCAG.....AGACGGGAAGTGGCCAAAACCGTCTTT 1215
      ::::::::::::::::::::
252 iSValLysLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeu 268
      ::::::::::::::::::::
1216 TCCTGGTCTGCTCTTCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
      ::::::::::::::::::::
269 ValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTy 285
      ::::::::::::::::::::
1266 GATTCTGAAGCTCACTCTTTTATAATCAGAAATGATCCCAATAGATGTGA 1315
      ::::::::::::::::::::
285 rLeuTyrrArgSerTyrrHisTyrrSerGluValAspThrSer.....M 299
      ::::::::::::::::::::
1316 TTTTTCAGCTTCTGTTGGTATTGGACTATATATTGCTATCAACATGCTTCA 1365
      ::::::::::::::::::::
299 etLeuHisPhe.....ValThrSerIleCysAlaHisLeuLeuAlaPhe 313
      ::::::::::::::::::::
1366 CTGAATTCCTGCATTAACCAATTCCTCTGCTATTTGGTCAGCAAAAGATT 1415
      ::::::::::::::::::::
314 ThrAsnSerCysValAsnProPheAlaLeuTyrrLeuLeuSerLysSerPh 330
      ::::::::::::::::::::
1416 CAAAACACTGCTTTAAGTCA...TGCTTATGCTGCTGCTGCTGCTGCTGCT 1462
      ::::::::::::::::::::
330 eArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGlyLeuMeta 347
      ::::::::::::::::::::
1463 AGAAAAACAGCTGCTTGGAGGAAAGACAGCTGCTGCTTAAAGTTCAAAGCT 1512
      ::::::::::::::::::::
347 snArgSerHisSerThrGlyArgSerThrThrCysMet..... 359
      ::::::::::::::::::::
1513 AATGATCAGCGATATGACAACTTCCTGCTTCCAGTAAT 1548
      ::::::::::::::::::::
360 .....ThrSerPheLysSerThrAsn 366
      ::::::::::::::::::::

```

seq_name: pir2:157682

seq_documentation_block:

bombesin/ GRP receptor - rat

C:Species: Rattus norvegicus (Norway rat)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999

C:Accession: 157682

R:Wada, E.; Way, J.; Shapira, H.; Battey, J.F.

Mol. Cell. Neurosci. 3, 446-460, 1992

A:Title: Two distinct bombesin receptor subtype subtypes in postnatal rat centra.

A:Reference number: 157682

A:Accession: 157682

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-384 <RES>

A:Cross-references: EMBL:X56661; NID:9509398; PIDN:CAA39988.1; PID:9509399

C:Superfamily: endothelin receptor B

alignment_scores:
 Quality: 496.50 Length: 366
 Ratio: 2.131 Gaps: 9
 Percent Similarity: 63.661 Percent Identity: 33.060

alignment_block:

US-09-931-157-2 x I57682

Align seg 1/1 to: I57682 from: 1 to: 384

```

484 CGCACATCTCCCTCCCGCTGGCAAGACCCATCGAGATCAAGAGAC 533
      ::::::::::::::::::::
25  GlnThrLeuSerProPro.....LysMetAspAsnTr 35
      ::::::::::::::::::::
534 TTTCAAA.....TACATCAACACGGTTGTGCTGCTGCTGCTGTGT 571
      ::::::::::::::::::::
35  pPheHisProGlyIleIleTyrrValIleProAlaValTyrrGlyLeuIle 52
      ::::::::::::::::::::
572 TCGTGTCTGGGATCATCTGGGAACCTCCACACTCTCTGAGAATATCTACAAG 621
      ::::::::::::::::::::
52  leValIleGlyLeuIleGlyAsnIleThrLeuIleLysIlePheCysThr 68
      ::::::::::::::::::::

```

[illegible]

```

1501 AAGTTCAAAGCTAATGATGATCAGGATATGACAACTTCGCTCCAGTAAT 1548
      360 .....ThrSerPheLysSerThrAsn 366
seq_name: pir2:B41007

seq_documentation_block:
bombersin receptor, neuromedin B-preferring - human
C:Species: Homo sapiens (man)
C:Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change
C:Accession: B41007
R:Corjay, M.H.; Dobrzanski, D.J.; Way, J.M.; Viallet, J.; Shapi
J. Biol. Chem. 266, 18771-18779, 1991
A:Title: Two distinct bombesin receptor subtypes are expressed
A:Reference number: A41007; MUID:92011639
A:Accession: B41007
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-390 <COR>
A:Cross-references: GB:M73482; NID:g189241; PIDN:AAA59939.1; PT
C:Genetics:
A:Gene: GDB:NMBR
A:Cross-references: GDB:128063; OMIM:162341
A:Map position: 6q21-6qter
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; transmembrane protein

alignment_scores:
      Quality: 491.00      Length: 367
      Ratio: 2.173      Gaps: 8
      Percent Similarity: 61.580      Percent Identity: 32.153

alignment_block:
US-09-931-157-2 x B41007 ..

Align seg 1/1 to: B41007 from: 1 to: 390

424 GGC GCGTTCGTGGCACCTGGGAGGTGCTCAAAGGAGACAGCAGCGCAGG 473
      :|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
3 SerLysSerLeuSerAsnLeuSerValThrThrGlyAlaAsnGluSerG1 19

474 ATCTCGGCCA.....CGCACATCTCCCTCCCTCCCGTGGCCAG 511
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
19 ySerValProGluGlyTyrGluAlaArgAspPheLeuProAlaSerAspGlyT 36

512 GACCCATCGAGATCAAGGAGACTTTCAAATACATCAACACGCTGTGTGTC 561
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
36 hrThrThrGluLeu.....ValIleArg 43

562 TGC.....CTGTGTTGCTGCTGGGATCATCG 590
      ||| |||||:|||||:|||||:|||||:|||||:|||||:|||||:
44 CysValIleProSerLeuTyrLeuLeuIleIleThrValGlyLeuLeuG1 60

591 GAACATCCACACTCTGAGATATATCTACAAGACAAGTGTCATCGAAACG 640
      |||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
60 yAsnIleMetLeuValIlystlePheIleThrAsnSerAlaMetArgSerV 77

641 GTCCCAATATCTGATCGCCAGCTTGGCTCTGGGAGACCTGCTGCACATC 690
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
77 a1ProAsnIlePheIleSerAsnLeuAlaIaGlyAspLeuLeuLeuLeu 93

691 GTCATGTGACATCCCTATCAATGTCTACAAGCTGCTGGCAGAGGACTGGCC 740
      :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
94 LeuThrCysValProValAspAlaSerArgTyrPhePheAspGluTyrPMe 110

741 APTTGGAGCTGAGATGTGAAGCTGGTGCCTTTTCATCAGAAAGACTCCG 790
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
110 tPheGlyLysValGlyCysLysLeuIleProValIleGlnLeuThrSerV 127

791 TGGGAATCACTGTGCTGAGTCTATGTGCTCTGAGTATTGACAGATATCGA 840
      ||||| |||||:|||||:|||||:|||||:|||||:|||||:|||||
127 alGlyValSerValPheThrLeuThrAlaLeuSerAlaAspArgTyrArg 143

```



```
841 GCTGTTGCTCTTGGAGTAGAATTAAGGAATTTGGGTTCCAAATGGAC 890
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
144 AlaIleValAsnProMetAspMetGlnThrSerGlyAlaLeuLeuArgTh 160
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
891 AGCAGTAGAAATGTTTTGATTTGGTGGTCTCTGTGGTCTGGCTGTCC 940
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
160 rCysValLysAlaMetGlyIlePValValSerValLeuLeuAlaValP 177
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
941 CTGAAGCCATAGCTTTTGATATAATTACGATG...GACTACAAAGGAAGT 987
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
177 roGluAlaValPheSerGluValAlaArgIleSerSerLeuLeuAspSer 193
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
988 TATCTCGAATCTGCTTGGTCTATCCCGTTTCAGAACACAGCTTTCATGCA 1037
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
194 SerPheThrAlaCysIleProTyPProGlnThrAspGluLeuHisProLy 210
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1038 GTTTTACAGACAGCAAAAGATTGGTGGTCTTCACTTCTTATTTCTGCT 1087
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 sIleHisSerValLeu.....IlePheLeuValTyPheLeuI 223
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1088 TGCATTGGCCATCAGTCACTGATTTTATATACACTAATGACCTGTGAATG 1137
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
223 IeProLeuAlaIleSerIleTyTyPThrHisIleAlaLysThrLeu 239
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1138 TTGAGAAAGAAAGTGGCATGCGATGCTTTAAATCATCATCAACGCA 1187
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
240 IleLysSerAlaHisAsnLeuProGlyGluTyPAsnGluHisThrLysLy 256
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1188 GAGACGGGAGTGGCCAAACCGTCTTTGCGTGGTCTTCTTGGCC... 1236
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
256 sGlnMetGluThrArgLysArgLeuAlaLysIleValLeuValPheValG 273
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1237 .....CTCTGCTGGCTTCCCTTCCACCTCAGCAGGATCTGAAG 1275
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
273 lYcysPheIlePheCysTyPTrpPheProAsnHisIleLeuTyPMetTyPArg 289
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1276 CTCACCTTTTATATCAGATGATCCCAATAGATGTAAGTCTTGGAGCTT 1325
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
290 SerPheAsnTyPAsnGluIleAspProSerLeuGlyHisMetIleValTh 306
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1326 TCTGTTG.....GTATTGGACTATATTGCTATCAACATGGCTTCACTGA 1369
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
306 rLeuValAlaArgValLeuSerPhe.....GlyA 316
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1370 ATTCTGCAATTAACCAATGCTCTGTATTGTTGGTGAAGAAAGATTCAAA 1419
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
316 snSerCysValAsnProPheAlaLeuTyPLeuLeuSerGluSerPheArg 332
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1420 AACTGCTTTAAGTCATGCTATGCTGCTGCCAGTCACTTTGAAGAAAA 1469
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
333 ArgHisPheAsnSerGlnLeuCysCysGlyArgLysSerTyPTrpGlnGluAr 349
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1470 A 1470
349 g 349
seq_name: pir2:JH0374
seq_documentation_block:
bombedesin receptor, neuromedin-B-preferring - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 16-Jul-1999
C:Accession: JH0374
R:Wada, E.; Way, J.; Shapira, H.; Kusano, K.; Lebacqz-Verheyden, A.M.; Coy, D.; Jensen, R.
Neuron 6, 421-430, 1991
A:Title: cDNA cloning, characterization, and brain region-specific expression of a neuro-
A:Reference number: JH0374; MUID:91158987
A:Accession: JH0374
A:Molecule type: mRNA
A:Residues: 1-390 <WAD>
A:Cross-references: GB:037058; NID:g1039481; PIDN:AAA79881.1; PID:g1039482
A:Experimental source: esophagus
```

C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F:42-65/Domain: transmembrane #status predicted <TD1>
F:80-99/Domain: transmembrane #status predicted <TD2>
F:118-139/Domain: transmembrane #status predicted <TD3>
F:157-177/Domain: transmembrane #status predicted <TD4>
F:212-235/Domain: transmembrane #status predicted <TD5>
F:267-287/Domain: transmembrane #status predicted <TD6>
F:300-327/Domain: transmembrane #status predicted <TD7>
F:8,71,192/Binding site: carbohydrate (Asn) (covalent) #status predicted

alignment_scores:

Quality: 489.50 Length: 363
Ratio: 2.205 Gaps: 8
Percent Similarity: 61.157 Percent Identity: 31.680

alignment_block:

US-09-931-157-2 x JH0374 ..

Align seg 1/1 to: JH0374 from: 1 to: 390

```
478 CGCCACGACGACCATCTCCCTCCCGTCCCAAGGACCCATCGAGATCAA 527
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
2 ProProArgSerLeu.....ProAsnLeuSerLeuProThrGluAlaSe 16
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
528 GGAGACT..... 534
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
16 rGluSerGluLeuGluProGluValTrpGluAsnAspPheLeuProAspS 33
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
535 .....TTCAAATACATCAACACGCTGTGTCTCTGC..... 564
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
33 erAspGlyThrThrAlaGluLeuValIleArgCysValIleProSerLeu 49
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
565 .....CTTGCTGTCTGGGATCATCGGGAACCTCCACACTTCTGAG 608
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
50 TyrLeuIleIleIleSerValGlyLeuLeuGlyAsnIleMetLeuVally 66
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
609 AATTATCTACAAGACAAGTGCATGCGAAACGGTCCCAATATCTTGTATCG 658
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
66 sIlePheLeuThrAsnSerThrMetArgSerValProAsnIlePheIleS 83
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
659 CCAGCTTGGCTCTGGGAGACCTGCTGCATCGTCATTCATGACATCCCTATC 708
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
83 erAsnLeuAlaLagIyAspLeuLeuLeuLeuLeuThrCysValProVal 99
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
709 AATGCTCTACAAGCTGCTGGCAGAGGACTGGCCATTTGGAGCTGAGATGTG 758
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 AspAlaSerArgTyPTrpPheAspGluTrpValPheGlyLysLeuGlyCy 116
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
759 TAAGCTGGTGGCTTTTCATACAGAAAGCCTCCGTGGGAATCACCTGTGCTGA 808
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
116 sLysLeuIleProAlaIleGlnLeuThrSerValGlyValSerValPheT 133
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
809 GTCTATGCTCTGAGTATTGACAGATATCGAGCTGTTGCTTCTTGGAGT 858
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
133 hrLeuThrAlaLeuSerAlaAspArgTyPArgAlaIleValAsnProMet 149
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
859 AGAATTAAAGGAATTGGGTTTCCAAATAGCAGACAGTAGAATTTGTTT 908
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
150 AspMetGlnThrSerGlyValValLeuTrpThrSerLeuLysAlaValG 166
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
909 GATTGGTGGTCTCTGTGTTCTGGCTGCTCCCTGAAGCCATAGGTTTTC 958
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
166 yIleTrpValValSerValLeuLeuAlaValProGluAlaValPheSerG 183
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
959 ATATATATTACGATG...GACTACAAAGGAAGTATTCTGCCAATCTGCTTG 1005
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
183 luValAlaArgIleGlySerSerAspAsnSerSerPheThrAlaCysIle 199
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1006 CTTTCATCCCTGTCAGAACAGACAGCTTTCATCCAGTTTTACAGACAGCAA 1055
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
200 ProTyPProGlnThrAspGluLeuHisProLysIleHisSerValLeu.. 215
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

57 yAsnIleThrLeuIleLysIlePheCysThrValLysSerMetArgAsnV 74

641 GTCCCAATATCTTGATGCGCAGCTTGGCTCTGGGAGACCTGCCTGCACATC 690
|||||:::|||||:::|||||:::|||||:::|||||:::
74 aProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLeuLeu 90

691 GTCATTGACATCCCATTCAATGTCTACAAGCTGCTGGCAGAGACTGGCC 740
::: |||||::: ::: |||||::: |||||::: |||||:::
91 IleThrCysAlaProValAspAlaSerArgTyLeuAlaAspArgTrpLe 107

741 ATTGGAGCTGAGATGTTGAAGCTGGTCTTCATACATACAGAAAGCCCTCCG 790
|||||:::|||||:::|||||:::|||||:::|||||:::
107 upheGlyargIleGlyCysLysLeuIleProPheIleGlnLeuThrServ 124

791 TGGGAATCACTGCTGCTGAGTCTATGTGCTCTGAGTATTGACAGATATCGA 840
|||||:::|||||:::|||||:::|||||:::|||||:::
124 aIcylValSerValPheThrLeuThrAlaLeuSerAlaAspArgTyLys 140

841 GCTGTGCTTCTTGAGTAGATAATAAGAATGGGTGCCAAAATGGAC 890
|||||::: |||||::: ::: |||||::: :::
141 AlaIleValArgProMetAspIleGlnAlaSerHisAlaLeuMetLysII 157

891 ACCAGTAGAAATGTTTTGATTGGGTGGTCTGCTGGTGTCTGGCTGCTCC 940
::: |||||::: |||||::: |||||::: |||||::: |||||:::
157 eCysLeuLysAlaAlaPheIleTrpIleSerMetLeuLeuAlaIle 174

941 CTGAAGCCATAGTTTGTGATATAATTACCATGCACTACAAGGAAGTAT 990
|||||::: |||||::: |||||::: |||||::: |||||:::
174 roGLuAlaValPheSerAspLeuHisPropheHisGluGluSerThrAsn 190

991 CTGCGAATC.....TGCTTGCTTCATCCCGTTCAAGAAGACACTTTCAT 1034
::: |||||::: |||||::: |||||::: |||||:::
191 GlnThrPheIleSerCysAlaProTyProHisSerAsnGluLeuHisPr 207

1035 GCAGTTTTACAGACAGCAAAGATGGTGCTGCTTCAGTTCTTCTATTTCT 1084
::: |||||::: |||||::: |||||::: |||||::: |||||:::
207 oLysIleHisSerMetAlaSer.....PheLeuValPheTyrv 220

1085 GCTTGCCATTGCCCATCACTGCATTTTATTATACACTAATAGACTGTGAA 1134
::: |||||::: |||||::: |||||::: |||||::: |||||:::
220 aIleProLeuSerIleIleSerValTyTyrrPheIleAlaLysAsn 236

1135 ATGTTGAGAAAGAAAGTGGCATGCGATGCTTTAAATGATCACCTAAA 1184
::: |||||::: |||||::: |||||::: |||||::: |||||:::
237 LeuIleGlnSerAlaTyrrAsnLeuProValGluGlyAsnIleHisValLy 253

1185 GCAG.....AGACGGGAAGTGGCCAAACCGCTTTTGCTCTGG 1222
|::: |||||::: |||||::: |||||::: |||||::: |||||:::
253 slysGlnIleGluSerArgLysArgLeuAlaLysThrValLeuValPhev 270

1223 TCCTTGCTTTGCCCTCTGCTGCTCCCTTCCACTCAGCAGGATCTCG 1272
|| ::|||:::|||||:::|||||::: |||||::: |||||:::
270 aIGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTyLeuTyrr 286

1273 AAGCTCACTCTTTTAATCAGAAATGATCCCAATAGATGTGAACCTTTGAG 1322
::: |||||::: |||||::: |||||::: |||||::: |||||:::
287 ArgSerTyrrHistyrrSerGluValAspThrSer.....MetLeuHi 300

1323 CTTTCTGTTGGTATTGGGACTATATTGGTATCAACATGGCTTCACTGAATT 1372
Phe ::|||::: |||||::: |||||::: |||||::: |||||:::
300 sPhe.....ValThrSerIleCysAlaArgLeuLeuAlaPheThrAsnS 315

1373 CQTGCATTAACCCAATTGCTGCTGATTTCGGTGAGCAAAAGATTCAAAAAC 1422
|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::
315 eCysValAsnProPheAlaLeuTyrrLeuLeuSerLysSerPheArgLys 331

1423 TGCTTTAAGTCATGCTTATGCTGGTGCCAG 1455
|||||::: |||||::: |||||::: |||||:::
332 GlnPheAsnThrGlnLeuLeuCys...CysGln 341

seq_name: pir2.A46632

seq_documentation_block:

bombesin-like peptide receptor 3 - human

C:Species: Homo sapiens (man)
C:Date: 22-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 16-Jul-1999
C:Accession: A46632; S42403
R:Path1, Z.; Cortaj, M.H.; Shapira, H.; Wada, E.; Benya, R.; Jensen, R.; Viallet, J.; Sa
J. Biol. Chem. 268, 5979-5984, 1993
A:Title: BRS-3: a novel bombesin receptor subtype selectively expressed in testis and lu
A:Reference number: A46632; MUID:93194908
A:Accession: A46632
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA: mRNA
A:Residues: 1-399 <FAT>
A:Cross-references: GB:L08893; NID:g291876; PIDN:AAA35604.1; PID:g291877
A:Experimental source: placenta
R:Gorbulev, V.; Akhundova, A.; Grzeschik, K.H.; Fahrenholz, F.
FEBS Lett. 340, 260-264, 1994
A:Title: Organization and chromosomal localization of the gene for the human bombesin re
A:Reference number: S42403; MUID:94178377
A:Accession: S42403
A:Molecule type: DNA

A:Residues: 1-399 <GOR>
A:Cross-references: GB:X76498; NID:g468753; PIDN:CAA54031.1; PID:g468754
C:Genetics:
A:Gene: GDB:BRS3
A:Cross-references: GDB:l138468
A:Map position: Xq26-Xq28
C:Superfamily: endothelin receptor B
C:Keywords: G protein-coupled receptor; transmembrane protein

alignment_scores:

Quality: 443.00 Length: 315
Ratio: 2.150 Gaps: 9
Percent Similarity: 65.397 Percent Identity: 33.016

alignment_block:

US-09-931-157-2 x A46632 ..

Align seg 1/1 to: A46632 from: 1 to: 399

562 TCCTCTGTTGGTGGTG.....CTGGGATCATCGG 590
48 CysAlaIleTyrIleThrTyrAlaValIleSerValGlyLeuGl 64
591 GAATCCACACTCTCAGAAATATCTACAAGAACAAAGTCATCGGAAACG 640
64 YAsnAlaIleLeuIleLysValPhePheLysThrLysSerMetGlnThr 81
641 GTCCCAATATCTTGTCCGACCTGGCTCTGGGACACCTGCTGCACATC 690
81 alProAsnIlePheIleThrSerLeuAlaPheGlyAspLeuLeuLeu 97
691 GTCATTGACATCCCTATCAATCTACAAGCTGCTGGCAGAGGACTGGCC 740
98 LeuThrCysValProValAspAlaThrHisTyrLeuAlaGluGlyTrpLe 114
741 ATTTGGAGCTGAGATCTGTAGCTGGTCTTCATACAGAAAGCTCCG 790
114 uPheGlyArgIleGlyCysValLeuSerPheIleArgLeuThrServ 131
791 TGGGAATCACTGCTGAGTCTATGCTGCTGAGTATTGACAGATATCGA 840
131 alGlyValSerValPheThrLeuThrIleLeuSerAlaAspArgTyrLys 147
841 GCTGTGCTCTTCTGGAGTAGAATTAAAGGAATTGGGTTCCAATGGAC 890
148 AlaValValLysProLeuGluArgGlnProSerAsnAlaIleLeuLysTh 164
891 AGCAGTAGAAATGTTTTCATTTGGTCTCTGTTGGTCTGCTGGCTGCC 940
164 rCysValLysAlaGlyCysValTrpIleValSerMetIlePheAlaLeuP 181

941 CTGAAGCCATAGGTTTTCATATATTACGATG.....GACTACAAAGGA 984
181 roGluAlaIlePheSerAsnValTyrThrPheArgAspProAsnLysAsn 197
985 AGTTATCTGCGAATCTGCTGCTTCATCCCGTTTCAGAAAGACAGCTTTCAT 1034
198 MetThrPheGluSerCysThrSerTyrProValSerLysLysLeuLeuGl 214
1035 CGAGTTTACAGACAGCAAAAGATTGGTGGCTGTTCAAGTTTCTATTCT 1084
214 nGluIleHisSerLeuLeu.....CysPheLeuValPheTyrI 227
1085 GCTTCCCATTCGCCCATCTACTCATTTTATACACTAATGACCTGTGAA 1134
227 leileProLeuSerIleIleSerValTyrSerLeuIleAlaArgThr 243
1135 ATGTTGAGAAAGAAAGTGGATGCAGATTGCTTTAAATGATCAC..... 1179
244 LeuTyrLysSerThrLeuAsnIleProThrGluGluGlnSerHisAlaAr 260
1180CTAAAGCAGACGCGGAAGTGGCCAAACCGCTCTTTTCCTGG 1222
260 gLysGlnIleGluSerArgLysArgIleAlaArgThrValLeuValLeuV 277
1223 TCCTTGTCTTGGCCCTCTGCTGGCTTCCCTTCACCTCAGCAGGATTCTG 1272
277 alAlaLeuPheAlaLeuCysTrpLeuProAsnHisLeu..... 289
1273 AAGCTCACTCTTTATAATCAG.....AATGATCCCAA 1304
290 ...LeuTyrLeuTyrHisSerPheThrSerGlnThrTyrValAspProSe 305
1305 TAGATGTGAACCTTTTGAGCTTCTCTGTTGGTATTGGACTATATGTTATCA 1354
305 r.....AlaMetHisPheIlePheThrIle.....PheSerArgV 317
1355 ACATGGCTTCACTGAATTCCTGCATTAAACCAATTGCTGCTGTTGGTG 1404
317 alLeuAlaPheSerAsnSerCysValAsnProPheAlaLeuTyrTrpLeu 333
1405 AGCAAAAGATTCAAAACCTGCTTTAAGTCATGCTTA...TGCTGC 1446
334 SerLysSerPheGlnLysHisPheLysAlaGlnLeuPheCysCys 348

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2002, 01:00:07 ; Search time 221.29 Seconds
(without alignments)
4774.141 Million cell updates/sec

Title: US-09-931-157-2

Perfect score: 4301

Sequence: 1 gagacattccggtggggac.....ctgggaaaaaaaaaaaaa 4301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA:*

- 1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4301	100.0	4301	4	US-08-121-446-3
2	1691.8	39.3	1873	4	US-08-910-864-13
3	395.8	9.2	4105	4	US-08-121-446-1
4	132.2	3.1	1700	5	PCT-US92-02091-1
5	114.6	2.7	246240	2	US-08-724-394A-20
6	114.6	2.7	246240	2	US-08-724-394A-21
7	114.6	2.7	246240	2	US-08-724-394A-22
8	108.8	2.5	1584	5	PCT-US92-02091-5
9	106.6	2.5	1205	3	US-09-120-772-1
10	106	2.5	1726	5	PCT-US92-02091-3
11	101.6	2.4	1352	5	PCT-US92-02091-7
12	98.6	2.3	1563	1	US-08-279-590A-1
13	98.6	2.3	1563	1	US-08-910-092-1
14	84.8	2.0	4156	1	US-08-465-687A-1
15	84.8	2.0	4156	3	US-09-030-970-1
16	84.8	2.0	4156	5	PCT-US94-11843-1
17	80.6	1.9	1443	3	US-08-959-381A-3
18	80.6	1.9	1626	3	US-08-959-381A-4
19	63.2	1.5	595	1	US-08-784-289-1
20	61.8	1.4	2624	1	US-08-232-144-3
21	61.8	1.4	4571	1	US-08-232-144-5
22	60.2	1.4	1152	3	US-09-045-186-1
23	60.2	1.4	1152	3	US-09-045-186-3
24	60.2	1.4	1633	4	US-09-119-788-1
25	59.6	1.4	1686	4	US-08-029-170-30
26	58.8	1.4	1506	1	US-07-937-609-13
27	58.8	1.4	1506	4	US-08-029-170-13

28	58.4	1.4	2144	4	US-08-876-798A-1	Sequence 1, Appli
29	58.4	1.4	2144	4	US-08-876-798A-3	Sequence 3, Appli
30	58.2	1.4	1330	3	US-08-147-592A-5	Sequence 5, Appli
31	58.2	1.4	1330	4	US-08-292-694A-5	Sequence 5, Appli
32	57.6	1.3	1280	1	US-08-192-288-1	Sequence 1, Appli
33	57.6	1.3	1280	2	US-08-687-355A-1	Sequence 1, Appli
34	57.6	1.3	1280	1	US-08-415-818-5	Sequence 5, Appli
35	57.6	1.3	2280	2	US-08-894-236-5	Sequence 5, Appli
36	57.6	1.3	2280	5	PCT-US96-01444-5	Sequence 5, Appli
37	56.8	1.3	837	4	US-08-998-416-288	Sequence 288, App
38	56.6	1.3	1610	3	US-08-889-108-7	Sequence 7, Appli
39	56.6	1.3	1610	5	PCT-US94-10358-7	Sequence 7, Appli
40	56.6	1.3	2160	4	US-08-188-275A-1	Sequence 1, Appli
41	56.6	1.3	2162	4	US-09-351-198-1	Sequence 1, Appli
42	56.6	1.3	2162	4	US-09-113-426-1	Sequence 1, Appli
43	55.8	1.3	665	2	US-08-883-795A-36	Sequence 36, Appli
44	55.8	1.3	19124	2	US-08-487-826B-13	Sequence 13, Appli
45	55.2	1.3	1200	2	US-08-687-355A-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-121-446-3
; Sequence 3, Application US/08121446
; Patent No. 6313276
; GENERAL INFORMATION:
; APPLICANT: IMURA, HIROO
; APPLICANT: NAKAO, KAZUWA
; APPLICANT: NAKANISHI, SHIGETADA
; TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: MORRISON & FOERSTER
; STREET: 755 Page Mill Road
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304-1018
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/121.446
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/911,684
; FILING DATE: 10-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: CIOTTI, THOMAS E.
; REGISTRATION NUMBER: 21,013
; REFERENCE/DOCKET NUMBER: 29900-20324.00
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 813-5600
; TELEFAX: (415) 494-0792
; TELEX: 706141
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4301 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..1566
; US-08-121-446-3

Query Match 100.0%; Score 4301; DB 4; Length 4301;

Best Local Similarity 100.0%; Pred. No. 0;				
Matches 4301; Conservative		Mismatches 0;	Indels 0;	Gaps
Qy	1	gaacattccggtggggactgcccagcccgacgaactggtatcctgagacactccc	60	
Db	1	GAGACATCCGGTGGGGACTCTGGCCACCCGACGACGTGGATCTCTTGAGAGACATCCC	60	
Qy	61	aggtaggcatttcccgcggtgggacgcttcgacagcagtggtggcagcggcccgctgg	120	
Db	61	AGGTAGGCATTTCCCGCGGTGGGACGCTTGCCAGAGCAGTGTGTGGCAGGCCCCCGCTGG	120	
Qy	121	aggatcaacacagtggtgaaacctgggaagaaatgtaacttggagtcgtggaactctga	180	
Db	121	AGGATCAACACAGTGGCTGAACACTGGGAAGAACTGGTACTCTGGAGTCTGGACATCTGA	180	
Qy	181	aacttgactgaaactcggaagcgcaccggaagcagcttctgagcaggtagcagcatg	240	
Db	181	AACCTTGACTGAAACTGCGAGCGCCACCGACGCTCTTGGAGCAGGTAGCAGCATG	240	
Qy	241	cagcgcctccaaagtctgcggaagcgccctgggtgcgtggcttctgctgcgagcgtg	300	
Db	241	CAGCGCCTCCAAAGTCTGTGCGGACGCGCCCTGGTGGCTGGTCTTGGCTGGCGCTG	300	
Qy	301	tcgcgatctggggagagagagagcttccgcgctgacaggccaactccgcttttgcga	360	
Db	301	TCGCGGATCTGGGAGAGGAGAGAGGCTTCCCGCTGACAGGGCCACTCCGCTTTTGCAA	360	
Qy	361	accgagagataatgacgcaccaccactaagaacttatggcccaagggttccaaagcagtg	420	
Db	361	ACCGCAGAGATAATGACGCCACCCACTAGACCTTATGGCCCCAAGGGTTCCACAGCCAGT	420	
Qy	421	ctggcgcggtcgcttgacactcggaaggtgacctaaagagagacagcagcgagatcttcg	480	
Db	421	CTGGCGCGGTCTGTGGCACCTGCGGAGGTGCTTAAAGNAGACAGGACGGCAGGATCTCCG	480	
Qy	481	ccagcaccactctcccctcccggtgccaaagaccactagagatcaaggagactttcaaa	540	
Db	481	CCAGCACCACTCTCCCTCCCTCCCGTGCCAAAGNACCATCGAGATCAAGGAGACTTTCAAA	540	
Qy	541	tacatcaaacaggttggtgctgctcgttgcgtgctgggatactcgggaactccaaca	600	
Db	541	TACATCAACAGGTTGGTCTCTGCCCTTGTGTGCTGTGGGATCATCGGAATCTCCACA	600	
Qy	601	cttctgagaattatclacaagaacaagtgcatacgcaaacggtcccaatatctctgacgcc	660	
Db	601	CTTCTGAGAAATATCTACAAGAACAAAGTGCATGCGAAGGGTCCCCAATATCTTGATCGCC	660	
Qy	661	agcttggctctgggagacctgtgcacatcgctcaattgacatccctatcaatgtctacaag	720	
Db	661	AGCTTTGGCTCTGGGAGACCTGCTGTCACATCGCTCATTGACATCCCTATCAATGTCTACAA	720	
Qy	721	ctgctgcagagacgtgccaattggagcgtgagatgtgaagctggtcccttccatacag	780	
Db	721	CTGCTGCGCAGAGACTGGCCATTTGGAGCTGAGATGTGTAAGCTGGTCCCTTCATACAG	780	
Qy	781	aaagccctccgtgggaatcacgtgctgagctctatgctgctgagtatgacagatatcga	840	
Db	781	AAAGCCTCCGTGGGAATCACTGCTGAGTCTATGTGCTCTGAGTATTTGACAGATATCGA	840	
Qy	841	gctgtgtcttcttgagtagaattaaaggaattgggggttccaaaatggacagcagtagaa	900	
Db	841	GCTGTGCTCTTTGGAGTAGAATTAAAGGAATTGGGTTTCCAAATGGACAGCAGTAGAA	900	
Qy	901	attgttttgattgggtggctctgctgtggttctggtctgcctggaagcataaggtttgat	960	
Db	901	ATTGTTTTGATTTGGGTGGTCTCTGTGGTCTGTGGTGTCCCTGGAAGCCATAGGTTTTGAT	960	
Qy	961	ataattacgatcactacaagaagatctatcgcaaatctgctgtctcatcccggttcag	1020	
Db	961	ATAATTACGATGGACTACAAGGAAGTTATCTGCGAATCTGCTTGTTCATCCCGTTCCAG	1020	
Qy	1021	aagacagctttcatgcagtttttacaagacagcaaaagattggtggctgttcagtttctat	1080	

Db	1021	AAAGACAGCTTTTCATGCAGTCTTTTACAAAGACAGCAAAAGATTTGGTGGCTGTTTCACGTTTCTAT	1081
Qy	1081	ttctctgtccattggccatccactgcattttttatcacactaatgacctgtgaaatgttg	1140
Db	1081	TTTCTGCTTGGCAATTGGCCATCACTGCATTTTATTATACACTAATGACCTGTGAATGTTG	1140
Qy	1141	agaaagaaaaagtgycatgcagattgcttttaaatgatcacctaaaagcagagacgggaagt	1200
Db	1141	AGAAAGAAAAAGTGGCATGCAGATTTGCTTTTAAATGATCACTAAAGCAGAGACGGGAAGTG	1200
Qy	1201	gccaaaaccgtcttttgcctgctctgtctcttgcctctgctgcttcccttccacct	1260
Db	1201	GCCAAAACCGCTTTTGGCTTGGTCCGTGTCTTTGGCCCTCTGCTGGGTTCCTCCCTACACCTC	1260
Qy	1261	agcagattctgaagctcactctttataatcagaatgatcccaatagatgtgaacttttg	1320
Db	1261	AGCAGATTCTGAAGCTCACTCTTTATAATCAGAATGATCCCAATAGATGTGAACTTTTG	1320
Qy	1321	agcttctgttggtattgactatatgtgatacaaatggcttcaactgaattcctgcatt	1380
Db	1321	AGCTTTCTGTGTATGACTATATTTGTGTATCAACATGGCTTTCACGTGAATTCCTGCATT	1380
Qy	1381	aaaccaattgctctgattgtgacaaagaattcaaaaactgctttaaagtcattgctta	1440
Db	1381	AAACCAATTTGCTCTATTTTGGTAGCAAAAGATTCAAAAACGTCTTAAAGTCATGCTTA	1440
Qy	1441	tactctgctgcagctcatttgaagaaaaaacagctcttgaggaaaaagcagtcgtgctta	1500
Db	1441	TGCTGCTGGTGGCGATCACTTTGAGAAAAAACAGTCTCTGGAGGAAAAAGCATGCTGCTTA	1500
Qy	1501	aagttcaaaagctaagatgacacggatagcaaacctccgttccagtaataataacagctca	1560
Db	1501	AAAGTCAAAGCTAAATGATCACGGATATGACAACTTCGTTCCAGTAATAAATACAGCTCA	1560
Qy	1561	tottgaaagaagaactattcactgtatttcattttttttatatattggaccgaagtcattaa	1620
Db	1561	TCCTGAAAGAAGAACTATTTCATCTGTATTCTTTCTTTTCTTATATTTGGACCGAAAGTCATTAA	1620
Qy	1621	aaacaaatgaacatttgcaaaacaaacaaaaaactatgtattgacagacacactat	1680
Db	1621	AACAAAATGAAACATTTGGCAAAACAAACAAAACATATGTATTTGCACAGCACACTAT	1680
Qy	1681	taaaatattaagtgttaatttttaaacctcacagctacatatgacattttatgagctgt	1740
Db	1681	TAAATATTAAAGTGTAAATATTATTAACTCACAGCTACATATGACATTTTATGAGCTGT	1740
Qy	1741	tiacggcatgaaagaagaatcagtggaattgaagaagcctcgtcgtggaagcactcaat	1800
Db	1741	TTACGGCATTGGAAAGAAATACATGGGAATTAAGAAGCCCTCGTCGTGGAAGCACTTAAT	1800
Qy	1801	ttttacagttagcacttcaacatagctcttaacaactccaggaattcacacaacact	1860
Db	1801	TTTTTACAGTTAGCACTTCAACATAGTCTTTAACAACTCCAGGATATTCACACAACACT	1860
Qy	1861	taggcttaaaatgagctcactcagaatttctctattctttctaaaaagagatttttta	1920
Db	1861	TAGGCTTAAAAATGAGCTCACTCAGAAATTTCTATTCTTTCTTAAAAAGAGATTTATTTTA	1920
Qy	1921	aatcaatgggactctgatataaaggagaataagtcactgtaaacagaacttttaagt	1980
Db	1921	AATCAATGGGACTCTGATATAAAGGAAGAAATAAGTCACCTGTAAACAGAACTTTTAAATG	1980
Qy	1981	aagcttaaatcactcaatttaaaatttaaaattcctttaaacaacttttcaatataat	2040
Db	1981	AGCTTTAAATTACTCAATTTAAATTTTAAATTCCTTTTAAACAACTTTTCAATATAT	2040
Qy	2041	tatcacactattatcagattgtaattagatgcaaatgagagagcagttttagttgttcgat	2100
Db	2041	TATCACACTATTATCAGATTGTAATTATAGATGCAATGAGAGAGCAGTTTATGTTGTCAT	2100
Qy	2101	ttttcggacactggaacaacttaaatgatacagaggaggagttaacagaagaagcaggtctg	2160
Db	2101	TTTTTCGGACACTGGAAACATTTAAATGATCAGAGGGAGTAAACAGAAGAGAAAGCGCTGT	2160

QY 2161 tttgaaatcattacacttttcaactagaagcccaaaccttcagcattctgcaatgtaac 2220
|||||
Db 2161 TTTTGAATCATTTACATTTTCACTAGAGCCCAACCTCAGCATTCGCAATATGTAAC 2220
QY 2221 caacatgctcacaaacagcagcagtgtacacagactgcaacatgctgcagctgaaatttaaaa 2280
|||||
Db 2221 CAACATGCTCACAACAGCAGCATGTAAACAGACTGGCAGCTGAAATTTAAAA 2280
QY 2281 tataacttttaaaagaaaattattacatcctttacattcagttlaagatcaaacctca 2340
|||||
Db 2281 TATAATACCTTTTAAAAAGAAAATTTTACATCTCTTTTACATTCAGTTAAAGATCAAACTCA 2340
QY 2341 caaagaaaatagaatgtttgaaagcctatcccaaaagacttttttgaactctgtcattca 2400
|||||
Db 2341 CAAAGAAAATACAATGTTTGAAGGCTATCCCAAAAGACTTTTGTGAATCTGCTATCA 2400
QY 2401 catacctgtgaagacaatactactctacaaattttttcagagattataaaatctcttttt 2460
|||||
Db 2401 CATACCCTGTGAAGACAATACTATCTACAATTTTTCAGGATTTAAAAATCTTCTTTTT 2460
QY 2461 tcaactcgtagcttaaacctctgttggttttgcctcgtgaatacttaccctacataca 2520
|||||
Db 2461 TCACATATCGTAGCTTAAACTCTGTTGGTTTGTCTCATCTGTAATACTTACCTTACATACA 2520
QY 2521 ctgcattgtagatgattaaatgaggcaggccctgtgctcatagctttacgatatgagat 2580
|||||
Db 2521 CTGCATGCTAGATGATTAATGAGGCGAGGCCCTGTGCTCATAGCTTTACGATGGAGAT 2580
QY 2581 gcagtgacctataataaagaactgtgaactgctgctggtgcagtgctccacatgcaaaagg 2640
|||||
Db 2581 GCAGTGACCTCATAATAAGACTGTGAACCTGGCTGGTGCACTGTCACATGACAAAGGG 2640
QY 2641 gcagtgacctctctcaccatgctgtggttgaataagttctcagcatatgtataat 2700
|||||
Db 2641 GCAGGTAGACCCTCTCACCCATGCTGTGGTTTAAATGGTTTCTAGCATATGTATAAT 2700
QY 2701 gctatagttaaaaatactatttttcaaaatcatcacagattagtagtattcaacagctacctg 2760
|||||
Db 2701 GCTATAGTTAAAATACTATTTTCAAAATCATACAGATTAGTACATTTAAACAGCTACCTG 2760
QY 2761 taaagcttactaaatttttctgtattatttttttaaatagccaatagaaaagttgcttg 2820
|||||
Db 2761 TAAAGCTTATTACTAAATTTTGTATTATTTTGTAAATAGCCAATAGAAAGTTTGTCTG 2820
QY 2821 acatggtgctttcttctcatcagagcgaactgcttttgagacgtgaagaacctctt 2880
|||||
Db 2821 ACATGTTGCTTTTCTTTCATCTAGAGCAAAACTGCTTTTGTAGACCTTGAAGAACCTCTT 2880
QY 2881 agctttgtgcttctcctgcctaattttttatatcttctaagcaaaagtgcttagatagctt 2940
|||||
Db 2881 AGCTTTGTGCTTCTCGCTCAATTTTATATCTTCTAAGCAAAAGTGCTTTAGGATAGCTT 2940
QY 2941 gggatgagatgctgtgaaagtagtatacaagagaaaaacggaagagagagaaatgaagtg 3000
|||||
Db 2941 GGGATCAGATGCTGTGTAAGAGTAGTACAAGAGAAAACGGAAGAGAGAGAAAATGAGGTG 3000
QY 3001 ggggtgaggaacccatgggacagattcccatctctagcctaaagcttcgctcaattgctt 3060
|||||
Db 3001 GGGTTGGAGAAACCCATGGGGACAGATTCCCATCTTACGCTTAACGTTCTGTCATTCGCT 3060
QY 3061 cgtcacatcaatgcaaaagctcgtgattttgttccagcaaaacacagtcgcaatgttctca 3120
|||||
Db 3061 CGTCACATCAATGCAAAAGGCTCGATTGTTTCCAGCAAAACACAGTGCATGCTTCTCA 3120
QY 3121 gagtgaacttcgaaaataaattgggccaagagctttaaactcgttcttaaaatatgcccac 3180
|||||
Db 3121 GAGTGACTTTCGAATAAATGGGCCCAAGAGCTTTTAACTCGGTCTTTAAAAATATGCCAA 3180
QY 3181 atttttacttgtttcttctttaaagctggccacatgcttgaaaataagctagtaag 3240
|||||
Db 3181 ATTTTTACTTGTGTTTCTTTTAAATAGGCTGGGCCCATGCTTGGAAATAAGCTAGTANTG 3240

QY 3241 ttgtttctgtcgaatttgaatgtgatggtacagtaaacacaaacccacaaatgtggcca 3300
|||||
Db 3241 TTGTTTCTGTCTCAATATGAATGTGATGTGTACAGTAACCAACCAACCAACCAATGTGGCCA 3300
QY 3301 gaaagaagagcgaataataatgaattacacaccataatggattctatttataaatcaccc 3360
|||||
Db 3301 GAAAGAAAGAGCAATTAATTAATTCACACACCATATGGATTCTATTTTATAAATCACCC 3360
QY 3361 acaaacctgttctttaaatttcacccaatcacctttttcagagggcctgtttatcatagaagt 3420
|||||
Db 3361 ACNAACTGTCTTCTTAAATTCATCCCAATCACCTTTTCAGAGGCTGTATCATAGAAGT 3420
QY 3421 catttagactctcaatttttaaatatttttgaaactcaactaatatttttcacagtttataa 3480
|||||
Db 3421 CATTTTAGACTCTCAATTTTAAATTAATTTTGAATCAGTAAATATTTTTCACAGTTTATTAA 3480
QY 3481 tatatttaattcttcttaaatatttagattatttttattaccatgactgaatttttaca 3540
|||||
Db 3481 TATATTAAATTTCTATTAAATTTTAGATTATTATTATTTACCATGTACTGAATTTTATACA 3540
QY 3541 tccctgataccctttctctccatgctcagtatcatgttctcttaattctcttgccaaattt 3600
|||||
Db 3541 TCCTGATACCTTTCTCTCTCCATGTCAGTATCATGTTCTCTAATTTATCTTGCCAAATTT 3600
QY 3601 tgaactacacacaaaagcatactctgcattattataataaaattgcatactcagtgctt 3660
|||||
Db 3601 TGAACCTACACACAAAAGCATACTTTGCATTATTATTAATAAAATTCATCTCAGTGGCTT 3660
QY 3661 tttaaaaaaattgttgattcctcaaaccttttaacatactgataagtaagaacaataataat 3720
|||||
Db 3661 TTTTAAAAAANTGTTTGATTCAAAACTTTTAACATACGTATAGTAAGTAAGAAACAATTTAAT 3720
QY 3721 tctttacatactcctcaaaccaagatagaaaaagtgctctatcgcttccaacttcaaaacatgt 3780
|||||
Db 3721 TTCTTTACATACCTCAAAACCAAGATAGAAAAGGTGCTATCGTTCAACTTCAAAACATGT 3780
QY 3781 ttcctagtaaggactttaatatagcaacacagacacaaaattattgttaacatagatgatta 3840
|||||
Db 3781 TTCTTAGTATTAAAGACTTTTAAATATAGCAACAGACAAAATTTATGTTTAACTAGATGCTTA 3840
QY 3841 cagctcaaaagatttataaaagattttaaacttatttctccttatttccactgctcaat 3900
|||||
Db 3841 CAGCTCAAAAGATTTTAAAAAGATTTTAACTTATTCTCCCTTATTATCCACTGCTAAT 3900
QY 3901 gtggatgtatgttcaaacacaccttttagtattgatagcttcatatagcccaaaaggaataca 3960
|||||
Db 3901 GTGGATGTATGTTCAAAACACCTTTTAGTATTGATAGCTTACATATGCCAAAGGAATACA 3960
QY 3961 gtttatagaacacatgggtatgctgtagtaactttataaaagtgtaataatacaaatgt 4020
|||||
Db 3961 GTTTATAGCAAAAACATGGGTATGCTGTAGCTAACTTTTATAAAAGTGTAATATAACAATGT 4020
QY 4021 aaaaaattatctctgggaggtatttttggttgcttaaaagtgctgaatagttactgattt 4080
|||||
Db 4021 AAAAAATTATATCTGGGAGGATTTTGTGTTGCCCTTAAAGTGGCTATAGTTACTGATTT 4080
QY 4081 ttattatgtaagcaaaacccaataaaaaatttaagtttttttaacaactaccttatttttc 4140
|||||
Db 4081 TTTATTATGTAGCAAAACCAATAAAAAATTTAAGTTTTTTTAAACAACCTTATTATTTC 4140
QY 4141 actgtacagacactaaattcctaataactaatgtattgttttaaaagaaataataatgtga 4200
|||||
Db 4141 ACTGTACAGACACTAAATTCATTAATAACTAATTTGTTTAAAGAAATAATAAATGTGA 4200
QY 4201 caagtgacattatttattgttaaatatacaattatcaagcaagtatgaagtatttcaatt 4260
|||||
Db 4201 CAAGTGCACATATTATTATGTTAAATATACAAATTTATCAAGCAAGTATGAAGTTATTCAATT 4260
QY 4261 aaatggcacattttctggtctggtgggaaaaaataaaaaa 4301
|||||
Db 4261 AAATGCCACATTTCTGGTCTCTGGGAAAAAATAAAAAA 4301

Qy	572	tcgtgctggggatcatcggggaattccacaactcttgagaattatctacaagaacaagtga	631
Db	292	ttctcgtgggcttctctgggcaaacatcatctgctggtagaattatctctcaccacagacacca	351
Qy	632	tgcgaaacggtcccaatatcttctgatcgccaagcttggtctctggagagacctctctcacatcg	691
Db	352	tgcggagtggtccccaacattcttcatcttTAACCTGGCTGCGGAGAGACTCTGCTGCTGC	411
Qy	692	tcattgacatcccttatcaatgtctacaagtctctggcagaggaactggccatttggagctg	751
Db	412	TGACCTCGGTCCTCCAGTGGATGCGCTCCGATACCTCTTTGATGAATGGGTCTTCGGCAAGC	471
Qy	752	agatgtgtaagctggtgctcttctatacagaagaagctctctggtggaatcaactgtgctgagtc	811
Db	472	TGGGCTGCAAACTCATCCAGCCATCCAGCTCACCCTCGGTGGGGGTTTCCGTGTTCACTC	531
Qy	812	tatgtgctctgagtattgacagatctgaactcttctctcttgagtagaataaaggaa	871
Db	532	TCAGGGCCCTCAGGCGTGCACAGGTACAGACTATCGTGGAAACCCCATGGACATCGACAGCT	591
Qy	872	ttaggggttccaaaatgacacagcagtagaanaattgttttgaatttgggtggtctctgtggttc	931
Db	592	CTGGTGTGGTGCTGTGGACACATTGAAGGCCGCTGGGCATCTGGGTGGTCTCTGTGCTGT	651
Qy	932	tggctgtccctcgaagccatagggttttggatata	963
Db	652	TGGCTGTCCCTGAGGCTGTGTCTTTTCGGAAGTA	683

[illegible]

Qy 932 tggctgtccctgaagccataggtttttgatata 963
|||||
pb 652 TGGCTGTCCCTGAGGCTGTGTTTTCGGAAGTA 683

: APPLICANT: TSUI, PING
 : APPLICANT: ELSHOURBAGY, NABIL
 : TITLE OF INVENTION: HUMAN BOMBESIN RECEPTOR SUBTYPE
 : TITLE OF INVENTION: 3
 : NUMBER OF SEQUENCES: 2
 : CORRESPONDENCE ADDRESS:
 :

```

: TITLE OF INVENTION: HUMAN BOMBESIN RECEPTOR SUBTYPE
:
: TITLE OF INVENTION: 3
:
: NUMBER OF SEQUENCES: 2
:
: CORRESPONDENCE ADDRESS:

```

; TITLE OF INVENTION: 3
 ; NUMBER OF SEQUENCES: 2
 ; CORRESPONDENCE ADDRESS:

CORRESPONDENCE ADDRESS:
ADDRESSEE: Ratner & Prestia

STREET: P.O. Box 980
CITY: Valley Forge

STATE: PA
COUNTRY: USA

ZIP: 19482
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible

```
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS
```

```

; SOFTWARE: FASTSEQ FOR WINDOWS VERSION 2.0
; CURRENT APPLICATION DATA:
;

```

; APPLICATION NUMBER: US/09/120,772
 ; FILING DATE: 22-JUL-1998
 ;

CLASSIFICATION: ;
PRIOR APPLICATION DATA: ;

APPLICATION NUMBER:
FILING DATE:

ATTORNEY/AGENT INFORMATION:
NAME: DROZDA, PAUL E.

NAME: Plescia, Paul F
REGISTRATION NUMBER: 23,031

REFERENCE/DOCKET NUMBER: GP-70505
TELECOMMUNICATION INFORMATION:

TELEPHONE: 610-407-0700
TELEFAX: 610-407-0700

TELEX: 846169
INFORMATION FOR SEO ID NO: 1:

SEQUENCE CHARACTERISTICS:
LENGTH: 1205 base pairs

```

; LENGTH: 1200 base pairs
; TYPE: nucleic acid

```

```
STRANDEDNESS: single
;
TOPOLOGY: linear
;
```


Db 694 GCAGGTACTGGCTGACAGATGGCTATTTGGCAGGATTGGCTGCAAACTGATCCCTTTA 753
Qy 776 tacagaaagccctccgtgggaatacactgtgctagtctagtctgtgctgtgattgacagat 835
Db 754 TACAGCTTACCTCTCTGGGGGTCTGTCTTCACACTACGCGGCTCTCGGCAGACAGAT 813
Qy 836 atcagagctgtgctctctggagtagaataaagggaattggggttccaaaaatggagacgag 895
Db 814 ACAAGCCATTGCTCGCGGCAATGGATATCCAGGCCCTCCCATGCCCTGATGAAGATCTGCC 873
Qy 896 tagaaattgttttgggtgggtgctgtggttctggttctggttccctgaagccatagtt 955
Db 874 TCAAGCCGCTTATCTGATCATCTCCATGCTGCTGCGCCATCCAGAGCGCGTGT 933
Qy 956 ttgatataattacgatgagactacaagggaagtattctg 993
Db 934 CTGACCTCCATCCCTTCATGAGGAAAGCACCACACAG 971

RESULT 11

PCT-US92-02091-7

; Sequence 7, Application PC/TUS9202091

; GENERAL INFORMATION:

; APPLICANT: Battley Jr., James F.

; APPLICANT: Corjay, Martha H.

; APPLICANT: Feldman, Richard I.

; APPLICANT: Harkins, Richard N.

; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Edwin P. Ching

; STREET: 1501 Harbor Bay Parkway

; CITY: Alameda

; STATE: CA

; COUNTRY: USA

; ZIP: 94501

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02091

; FILING DATE: 19920313

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/426,150

; FILING DATE: 24-OCT-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/533,659

; FILING DATE: 05-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Ching, Edwin P.

; REGISTRATION NUMBER: 34090

; REFERENCE/DOCKET NUMBER: A-0092C

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-266-7476

; TELEFAX: 415-266-7400

; INFORMATION FOR SEQ ID NO: 7:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1352 base pairs

; TYPE: NUCLEIC ACID

; STRANDEDNESS: double

; TOPOLOGY: linear

; MOLECULE TYPE: cDNA to mRNA

; HYPOTHETICAL: NO

; ORIGINAL SOURCE:

; ORGANISM: Homo sapiens

; CELL TYPE: Small cell lung carcinoma

; CELL LINE: NCI-H345

; FEATURE:

; NAME/KEY: CDS

; LOCATION: 140..1312

RESULT 12

PCT-US92-02091-7

Query Match 2.4%; Score 101.6; DB 5; Length 1352;
Best Local Similarity 47.6%; Pred. No. 1e-14;
Matches 444; Conservative 0; Mismatches 464; Indels 24; Gaps 4;

Qy 545 tcaaacacggttgtgctcgtcctgtgtgctgctggggtatcatcgggaaatccacacttc 604
Db 273 TGATCCCGTCCCTCCTACCTGCTCATCATCACCCTGGCTGGCTGGGCAACATCATGCTGG 332
Qy 605 tgagaattatctacaagaacaagtgcgcgaacacgttcccaatatcttgatgcgcagct 664
Db 333 TGAAGATCTTCATCAACAACAGGCCCATGAGGAGCGTCCCAACATCTTCATCTCTAACCC 392
Qy 665 tggctctgggagacacctgcacatcgctcattgacatccatcatcaatgtctacaagtgc 724
Db 393 TGGCGCGCGGGACTTGTGCTGCTCCTCACCCTGCTCGCGGTGGAGCGCTCGCGGTACT 452
Qy 725 tggcagaggaactggccatttggagctgagatgtgtaagctggtgccttttcatacagaaag 784
Db 453 TCTTCGACGAGTGGATGTTTGGCAAGGTGGCTGCAAACTGATCCCTGTCTATCCAGCTCA 512
Qy 785 cctccgtgggaatcaactgtgctgagctctatgtgctctgagttatgcaagatatcgagctg 844
Db 513 CTTCCGTGGGGGTTCCTCGTGTTCACCTCCTCAGCGCCGACAGGTACAGAGCCCA 572
Qy 845 ttgcttcttggagtagaattaaaaggaaattgggggttccaaaaatggacagcagtagaaattg 904
Db 573 TCGTTAAACCCCATGGACATGCAGACGCTCAGGGGCATGCTGCGGAGCTGTGTGAAGGCCA 632
Qy 905 ttttgattgggtggtctctgtggttctggtcctgcctgaagccataggttttgatataa 964
Db 633 TGGGTATCTGGGTGGTCTCCGTGTGCTGGCAGTTCCTCCGAGC----GGTGTTCAGAA 688
Qy 965 ttacagtgagactacaaaggaaagtattctcgaaatgctgcttgcctccctccggttcagaaga 1024
Db 689 GTGGCTCGCATCAGTAGCTTGGATAATAGCAGCTTCACAGCATGTATCCCATACCTCAA 748
Qy 1025 cagcttctcagctgttttacagacagcaaaagattgggtggtgctcagttctctattct 1084
Db 749 ACAGATGAATTACATCCAAAGATTTCATTTCAGTGT-----CAFTTCTTGGTCTATTTC 803
Qy 1085 gcttgccttggccatcactgcatcttttttacacataatgacctgtgaaattgtgagaa 1144
Db 804 TCATACCACCTTCTATTATTAGCATTTATTATATATATATATATATATATATATATATAT 863
Qy 1145 agaaaagtggcagcagattgttttaaatgatacaccctaaagcagagacgggaagtggcca 1204
Db 864 GCGCACACAATCTCTCTGGAGAATACATGAACATACCAAAAAACAGATGGAACACACGGA 923
Qy 1205 aaaccgcttcttgcctggctccttcttgcctcctctgctggttcccttccactcagca 1264
Db 924 AACCGCTGGCTAAATTTGCTGTGTCTTGTGGGCTGTTCAT-----CTTCTGTT 974
Qy 1265 ggattctgaagctcactctttataatcagaatgatcccaatagatgtgaacttttgagct 1324
Db 975 GGTTCACCAACACATCCTTTACATGCTATCGGTCTTTCAACTATATATATATATATATAT 1034
Qy 1325 ttct-----gttgggtattgggcatatttggatatcaacatggcttccactgaatctctgca 1378
Db 1035 CTCTAGGCCACATGATTTGCACCTTAGTTGCCGGGTCTCAGTTCAGTTCAGTTCAGTTCAGT 1094
Qy 1379 ttaaccccaattgctctgtatttggtagcaaaagattcaaaaactctttaaagctatgct 1438
Db 1095 TCAACCCATTTGCTCTTTACCTACTCAGTGAAGGCTTCAGAGGAGCATTTTCAACAGCCAAC 1154
Qy 1439 tatgctgctggtggcagtcatttgaagaaaa 1470
Db 1155 TCTGCTGTGGGAGGAAGTCTCTATCAAGAGAGA 1186

```

QY      938  tcctgaagccataggtttgatat 962
      || |||| | | |||
Db      727  CTCGAGAAGCTGTGTTCTCAGATT 751

RESULT 13
US-08-910-092-1
: Sequence 1, Application US/08910092
: Patent No. 5814463
: GENERAL INFORMATION:
: APPLICANT: Spindel, Eliot R.
: APPLICANT: Nagalla, Srinivasa
: APPLICANT: Barry, Brenda
: TITLE OF INVENTION: NUCLEIC ACIDS ENCODING RECEPTORS
: TITLE OF INVENTION: FOR BOMBESIN-LIKE PEPTIDES
: NUMBER OF SEQUENCES: 10
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 225 Franklin Street
: CITY: Boston
: STATE: MA
: COUNTRY: US
: ZIP: 02110-2804
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: Windows95
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/910,092
: FILING DATE: 12-AUG-1997
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/279,590
: FILING DATE: 22-JUL-1994
: ATTORNEY/AGENT INFORMATION:
: NAME: Tsao, Y. Rocky
: REGISTRATION NUMBER: 34,053
: REFERENCE/DOCKET NUMBER: 00537/098002
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617/542-5070
: TELEFAX: 617/542-8906
: TELEX: 200154
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1563 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: FEATURE:
: NAME/KEY: Coding Sequence
: LOCATION: 192...1319
: US-08-910-092-1

```

	Query Match	2.38;	Score 98.6;	DB 1;	Length 1563;
	Best Local Similarity 53.5%;	Pred. No. 5.7e-14;			
	Matches 206;	Conservative 0;	Mismatches 179;	Indels 0;	Gaps 0;
y	578	tgggatacatcggaactccacactctcgagaattatctacaagaacaagtgcgatgcgaa	637		
b	367	TCGGTATCCTTGGAAACACAATACTTATAAAGATATTTTAAAAATCAAGTCAATGCAGA	426		
y	638	acggtgccaaatatctgataccagctggctctgggagacactctgcacatgcgcatcg	697		
b	427	CTGTTCCCTAATATTTTCATACACGCGCTGGCTTTGGAGATCTTCTTCTACTGCTGACCT	486		
y	698	acatccctctacaatgctctacaagctgctggcagaggactggccatttggagactgagatgt	757		
b	487	GGGTGCCAGTGGACGCATCTCGGTATATTGTGGACACGCTGGATGTTTGGAAAGACTGGCT	546		
y	758	gtaagctgggtgcctcttcatacagaagaagcctccgtggggaatcactctgtagtctatgtg	817		

Tue Oct 1 09:27:51 2002

Qy 729 agaggaactggccatttgagctgagatggtgaagctggtgcctttccatacacagaaagcctc 788
Db 2750 CAAGAAAGTGGCTGGTGGAGGACTTCTCCTGCAAGATCGGCCCTATATAGAGGTGCGCTTC 2809
Qy 789 cgtgggaatcactgtgctgagctctatgtgctctgagttatgacagatatcgagctg 844
Db 2810 TCTGGGAGTCACCACTTTACCTTATGTGCTCTGTGCATAGACCCGCTTCCGTGCTG 2865

Search completed: October 1, 2002, 01:16:39
Job time: 28877 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2002, 01:01:38 ; Search time 1153.8 seconds
(without alignments)
6400.112 Million cell updates/sec

Title: US-09-931-157-2
Perfect score: 4301
Sequence: 1 gagacattccgtg9gggac.....ctgggaaaaaaaaaaaaa 4301

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues
Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :				N_Geneseq_032802.*	
1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*				
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*				
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*				
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*				
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*				
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*				
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*				
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*				
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*				
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*				
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*				
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*				
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*				
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*				
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*				
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*				
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*				
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*				
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*				
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*				
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*				
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*				
23:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*				
24:	/SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*				

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4297.8	99.9	4301	AAQ34584	ETb receptor gene.
2	4284.4	99.6	4286	AAF21284	Human low adenosin
3	4284.4	99.6	4286	AAA35162	Human adenosine re
4	4284.4	99.6	13611	AAF21288	Human low adenosin
5	4284.4	99.6	13612	AAA35166	Human adenosine re
6	1691.8	39.3	1873	AAV17875	Homo sapiens endot
7	1690.8	39.3	1872	AAF21283	Human low adenosin
8	1690.8	39.3	1872	AAF21283	Human adenosine re
9	1495.4	34.8	1719	AAF21285	Human low adenosin

10	1495.4	34.8	1720	21	AAA35163	Human adenosine re
11	1389	32.3	1632	24	AAQ24966	Human G-protein co
12	1322.6	30.8	1329	23	ABI97988	Non-endogenous hum
13	1220.4	28.4	1578	21	AAF21287	Human low adenosin
14	1220.4	28.4	1578	21	AAA35165	Human adenosine re
15	1099.2	25.6	1958	24	ABI99321	Mouse ischaemic co
16	1091	25.4	1965	13	AAQ25892	Sequence euclidean
17	1066	24.8	1406	15	AAQ51922	Bovine ET receptor
18	395.8	9.2	1284	15	AAQ63209	Human endothelin r
19	395.8	9.2	1868	21	AAF20903	Human low adenosin
20	395.8	9.2	1868	21	AAF20915	Human endothelin r
21	395.8	9.2	1868	21	AAA34781	Human adenosine re
22	395.8	9.2	1868	21	AAA34793	Human adenosine re
23	395.8	9.2	2008	21	AAF20904	Human low adenosin
24	395.8	9.2	2008	21	AAA34782	Human adenosine re
25	395.8	9.2	4105	14	AAQ34583	ETa receptor gene.
26	395.8	9.2	5036	21	AAF21447	Human endothelin r
27	395.8	9.2	117609	21	AAF21435	Human receptor-rel
28	395.4	9.2	4105	21	AAA38341	Human endothelin r
29	394.2	9.2	1310	21	AAF20902	Human endothelin r
30	394.2	9.2	1310	21	AAF20914	Human ELAM-1 polyn
31	394.2	9.2	1310	21	AAA34780	Human adenosine re
32	394.2	9.2	1310	21	AAA34792	Human ELAM-1 polyn
33	394.2	9.2	146981	21	AAF21442	Human ELAM-1 polyn
34	392.6	9.1	1284	23	ABI97987	Non-endogenous hum
35	272.4	6.3	292	14	AAQ59707	Human brain Expres
36	182.6	4.2	230	21	AAQ04336	Human secreted pro
37	149.4	3.5	967	23	AAQ64944	DNA encoding novel
38	148.8	3.5	1889	22	AAH18585	Human CDNA sequenc
39	147.2	3.4	8942	22	AAF29338	Atopy related gene
40	147.2	3.4	33030	22	AAF29337	Atopy related gene
41	144.6	3.4	446	23	AAQ76292	DNA encoding novel
42	143	3.3	8812	22	AAH57526	Human pancreas cel
43	138.2	3.2	1527	23	AAQ70588	DNA encoding novel
44	136.8	3.2	114955	20	AAQ53491	Human adenosine A1
45	133.8	3.1	1464	22	ABA09678	Human bone marrow

ALIGNMENTS

RESULT 1
AAQ34584
ID AAQ34584 standard; DNA; 4301 BP.

AC AAQ34584;

DT 11-MAY-1993 (first entry)

XX ETb receptor gene.

XX Human; ETa; ETb; endothelin; receptor; transmembrane domain; N tail; extracellular; cytoplasmic; C tail; post translational; bovine; modification; ET-1 receptor; antagonist; circulatory system; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 238..1566

FT sig_peptide /*tag= a

FT mat_peptide /*tag= b

FT polyA_signal 2595..2600

FT polyA_signal 3134..3139

FT polyA_signal 3638..3643

FT polyA_signal 4101..4106

FT polyA_signal /*tag= g

FT polyA_signal 4258..4263

FT misc_feature /*tag= h
 FT 1909..1913
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= i
 FT 1997..2001
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= j
 FT 2119..2123
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= k
 FT 2273..2277
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= l
 FT 2745..2749
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= m
 FT 3346..3350
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= n
 FT /function= Related with mRNA instability
 FT 3484..3488
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= o
 FT 3495..3499
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= p
 FT /function= Related with mRNA instability
 FT 3632..3636
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= q
 FT 3852..3856
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= r
 FT /function= Related with mRNA instability
 FT 4108..4112
 FT /function= Related with mRNA instability
 FT misc_feature /*tag= s
 FT 4213..4217
 FT /function= Related with mRNA instability
 FT /*tag= t
 FT /function= Related with mRNA instability
 XX
 PN EP522868-A.
 XX
 XX 13-JAN-1993.
 PD
 XX 10-JUL-1992; 92EP-0306347.
 XX
 XX 12-JUL-1991; 91JP-0172828.
 XX
 PA (SHIO) SHIONOGI SEIYAKU KK.
 XX
 XX Imura H, Nakanishi S, Nakao K;
 PI WPI; 1993-010677/02.
 XX P-PSDB; AAR30886.
 DR
 XX Human Eta and Etb endothelin receptors - for measuring endothelin
 PT and screening for endothelin antagonists
 XX
 PS Claim 12; Fig 2; 39pp; English.
 XX
 CC The sequences given in AA034583-84 encode the human Eta and Etb
 CC endothelin receptors respectively. Eta is a 427 amino acid protein
 CC with a molecular weight of 48,726. Etb comprises 442 amino acids and
 CC has a molecular weight of 49,629. Eta has a higher affinity for
 CC endothelin (ET)-1 and ET-2, whereas Etb has no selectivity for ET-1,
 CC ET-2 or ET-3. The receptors each contain seven transmembrane domains
 CC and have an extracellular N tail and a cytoplasmic C tail. There are
 CC several potential sites for post translational modification, these
 CC sites are identical to those of bovine ET-1 receptor. Eta cDNA is
 CC 91.2% homologous to bovine ET-1 receptor cDNA and Etb cDNA is 61.1%
 CC homologous to that of bovine Eta-receptor. The receptor proteins are
 CC useful as reagents for measuring the amount of Et or screening for
 CC antagonists of the Et receptor when studying the circulatory system.
 XX
 XX Sequence 4301 BP; 1342 A; 830 C; 815 G; 1314 T; 0 other;

Query Match 99.9%; Score 4297.8; DB 14; Length 4301;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 4299; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 Qy 1 gagacattccggtggggactctgtgccagcccgagcaacgctgagatcctgagagcactccc 60
 Db 1 gagacattccggtggggactctgtgccagcccgagcaacgctgagatcctgagagcactccc 60
 Qy 61 aggtaggcatttgcgccggtgggacgccttgcacagacagctgtgtggaagcccccgtagg 120
 Db 61 aggtaggcatttgcgccggtgggacgccttgcacagacagctgtgtggaagcccccgtagg 120
 Qy 121 aggtatcaacacagtggtggaacactgggaagaaactggtactctgagctggaacatctga 180
 Db 121 aggtatcaacacagtggtggaacactgggaagaaactggtactctgagctggaacatctga 180
 Qy 181 aacttggctctgaaactcggagcggccacccgagccctcttggagcaggtagcagcatg 240
 Db 181 aacttggctctgaaactcggagcggccacccgagccctcttggagcaggtagcagcatg 240
 Qy 241 cagcgcctccaaagtctgtgcgacgcgcctctgtgctggttcttgcctgcgcgcctg 300
 Db 241 cagcgcctccaaagtctgtgcgacgcgcctctgtgctggttcttgcctgcgcgcctg 300
 Qy 301 tcgcggatctgggagagagagaggtctccgcctgcagagccactccgctttttgcaa 360
 Db 301 tcgcggatctgggagagagagaggtctccgcctgcagagccactccgctttttgcaa 360
 Qy 361 accgcagagataatgacgcccccactaagacacttatggcccaagggttccaaacgcagt 420
 Db 361 accgcagagataatgacgcccccactaagacacttatggcccaagggttccaaacgcagt 420
 Qy 421 ctgcgcggtctgttgcacctgcggaggtgcctaaaggagacagacgcgcagagatctcg 480
 Db 421 ctgcgcggtctgttgcacctgcggaggtgcctaaaggagacagacgcgcagagatctcg 480
 Qy 481 ccacgcacatctccctcccccgtgccaaagaccatcgagatcaagagagacttttcaaa 540
 Db 481 ccacgcacatctccctcccccgtgccaaagaccatcgagatcaagagagacttttcaaa 540
 Qy 541 tacatcaaacacggttgcctgcctgtgttcgtgctggggtatcatcggaactccaca 600
 Db 541 tacatcaaacacggttgcctgcctgtgttcgtgctggggtatcatcggaactccaca 600
 Qy 601 ctctgagaattatctacaagaacaagtgcatacgcaaacggttcccaatatcttgatcgcc 660
 Db 601 ctctgagaattatctacaagaacaagtgcatacgcaaacggttcccaatatcttgatcgcc 660
 Qy 661 agcttggctctggagacactgtgcacatcgctcactatgcattgacatccctcaatctacaag 720
 Db 661 agcttggctctggagacactgtgcacatcgctcactatgcattgacatccctcaatctacaag 720
 Qy 721 ctgctggcagagagactggccatttgagctgagatggttaagctggtgcttcttcacag 780
 Db 721 ctgctggcagagagactggccatttgagctgagatggttaagctggtgcttcttcacag 780
 Qy 781 aaagcctccgtgggaatcactgtgctgagctatgctgctgagattgacagatatacga 840
 Db 781 aaagcctccgtgggaatcactgtgctgagctatgctgctgagattgacagatatacga 840
 Qy 841 gctcttctcttgagatgagataaaggaattggggttccaaaatgacacagcagataga 900
 Db 841 gctcttctcttgagatgagataaaggaattggggttccaaaatgacacagcagataga 900
 Qy 901 attgtttgatttgggtgctctgtggttctgctgctgctccctgaagccatagttttgat 960
 Db 901 attgtttgatttgggtgctctgtggttctgctgctgctccctgaagccatagttttgat 960
 Qy 961 ataattacgattgactacaagaaggaatttatctgcgaatctgctgttcttccatcccgctcag 1020
 Db 961 ataattacgattgactacaagaaggaatttatctgcgaatctgctgttcttccatcccgctcag 1020

QY 1021 aagacagcttccagtcagttttacaagacagcaaaagattgggtggttcagtttctat 1080
DB 1021 aagacagcttccagtcagttttacaagacagcaaaagattgggtggttcagtttctat 1080
QY 1081 ttctgttgcattggccatcaactgcaattttttatatacactaatagcaactgtgaaatgtg 1140
DB 1081 ttctgttgcattggccatcaactgcaattttttatatacactaatagcaactgtgaaatgtg 1140
QY 1141 agaaagaaagtggcatgcagattgctttaaatgatacactaaagcagagacgggaagtg 1200
DB 1141 agaaagaaagtggcatgcagattgctttaaatgatacactaaagcagagacgggaagtg 1200
QY 1201 gccaaaaccgtcttttggcctgcttggcttcttggcctctgctggcttcccttcaactc 1260
DB 1201 gccaaaaccgtcttttggcctgcttggcttcttggcctctgctggcttcccttcaactc 1260
QY 1261 agcagatttgaagctcaactctttataatcagaatgatacccaatagatggaacttttg 1320
DB 1261 agcagatttgaagctcaactctttataatcagaatgatacccaatagatggaacttttg 1320
QY 1321 agctttctgtggtattggactatattggtatcaacatggcttcaactgaaatccctgcatt 1380
DB 1321 agctttctgtggtattggactatattggtatcaacatggcttcaactgaaatccctgcatt 1380
QY 1381 aocccaatgctctgtatttggtagcaaaagattccaaaactgctttaaagtcatgtta 1440
DB 1381 aocccaatgctctgtatttggtagcaaaagattccaaaactgctttaaagtcatgtta 1440
QY 1441 tgetgtgtggccagtcatttgaagaaaaaacagtccttggaggaagaaagcagtcgtgtta 1500
DB 1441 tgetgtgtggccagtcatttgaagaaaaaacagtccttggaggaagaaagcagtcgtgtta 1500
QY 1501 aagttcaaaagctaatgatcacggatagacaacttccgttccagtaataaatcacagctca 1560
DB 1501 aagttcaaaagctaatgatcacggatagacaacttccgttccagtaataaatcacagctca 1560
QY 1561 tcttgaagaagaactatccaactgtatttcaatttcttatttcttatttggaccggaagtcataa 1620
DB 1561 tcttgaagaagaactatccaactgtatttcaatttcttatttcttatttggaccggaagtcataa 1620
QY 1621 aacaaaatgaacatttgcacaaacaaacaaaacacatgtatttgcacagcacacat 1680
DB 1621 aacaaaatgaacatttgcacaaacaaacaaaacacatgtatttgcacagcacacat 1680
QY 1681 taaaattaaagttaatttttaacactcacagctacatatgacatttttatgagctgt 1740
DB 1681 taaaattaaagttaatttttaacactcacagctacatatgacatttttatgagctgt 1740
QY 1741 ttacggcatggaagaaatacagtcgggaattaaagaagcctcgtcgtgaaagcacttaat 1800
DB 1741 ttacggcatggaagaaatacagtcgggaattaaagaagcctcgtcgtgaaagcacttaat 1800
QY 1801 tttttacagttagcacttcaactagctttaaacaacttccagatattcacacaact 1860
DB 1801 tttttacagttagcacttcaactagctttaaacaacttccagatattcacacaact 1860
QY 1861 taggcttaaaaatgagctcactcagaatttctatttcttcttaaaaagagattttattta 1920
DB 1861 taggcttaaaaatgagctcactcagaatttctatttcttcttcttaaaaagagattttattta 1920
QY 1921 aatcaatgggactctgatatataaaggaaataaagtcactgttaaaacagaaacttttaaatg 1980
DB 1921 aatcaatgggactctgatatataaaggaaataaagtcactgttaaaacagaaacttttaaatg 1980
QY 1981 aagcttaaatctactcaatttaaaattttaaaatccctttaaacaacttttcaattaatat 2040
DB 1981 aagcttaaatctactcaatttaaaattttaaaatccctttaaacaacttttcaattaatat 2040
QY 2041 tatcacactattatcagattgtaattagatgcaaatgagagagcagtttagttgtgcat 2100
DB 2041 tatcacactattatcagattgtaattagatgcaaatgagagagcagtttagttgtgcat 2100

QY 2101 ttttcggacactggaacatttaaatgatcaggaggagtaacagaaagcagcaggtgtg 2160
DB 2101 ttttcggacactggaacatttaaatgatcaggaggagtaacagaaagcagcaggtgtg 2160
QY 2161 ttttgaaaaatcattacactttcactagaagcccaaacctcagcattctgcaaatgttaac 2220
DB 2161 ttttgaaaaatcattacactttcactagaagcccaaacctcagcattctgcaaatgttaac 2220
QY 2221 caacatgtccacaaacagcagcatgttaacagactggcacatgtgccagctgaaatttaaaa 2280
DB 2221 caacatgtccacaaacagcagcatgttaacagactggcacatgtgccagctgaaatttaaaa 2280
QY 2281 tataactttttaaagaataatttataatcctttcatttccattcagtaagatacaactca 2340
DB 2281 tataactttttaaagaataatttataatcctttcatttccattcagtaagatacaactca 2340
QY 2341 caaagagaataagaatttgaagagctatcccaaaagactttttgaatctgcaatca 2400
DB 2341 caaagagaataagaatttgaagagctatcccaaaagactttttgaatctgcaatca 2400
QY 2401 cataccctgtgaagacaataactatctacaatttttccagattattaaaaacttctcttt 2460
DB 2401 cataccctgtgaagacaataactatctacaatttttccagattattaaaaacttctcttt 2460
QY 2461 tcactatcgtagcttaaaactctgtttggttttgttcactgttaataacttaactacataca 2520
DB 2461 tcactatcgtagcttaaaactctgtttggttttgttcactgttaataacttaactacataca 2520
QY 2521 ctgcatgtagatgattaaatgagggcagccctgtgctcatagcttttacgatggagagat 2580
DB 2521 ctgcatgtagatgattaaatgagggcagccctgtgctcatagcttttacgatggagagat 2580
QY 2581 gccagtgacctcacaataaaagactgtgaactgctgtgtgagctgtccacatgcacaaaggg 2640
DB 2581 gccagtgacctcacaataaaagactgtgaactgctgtgtgagctgtccacatgcacaaaggg 2640
QY 2641 gcagtgagcaccctctctcaccatgctgtggtttaaaaatggtttctcagcatatgataat 2700
DB 2641 gcagtgagcaccctctctcaccatgctgtggtttaaaaatggtttctcagcatatgataat 2700
QY 2701 gctatagttaaaaactatattttcaaaatcatagattagatagcatattaaacagctacctg 2760
DB 2701 gctatagttaaaaactatattttcaaaatcatagattagatagcatattaaacagctacctg 2760
QY 2761 taaagcttattactaaatttttgtattattttttaaataagccaatagaaaagtttgcttg 2820
DB 2761 taaagcttattactaaatttttgtattattttttaaataagccaatagaaaagtttgcttg 2820
QY 2821 acatggtgcttttcttcacatagagcaaaactgctttttgagaccgttagaacctctt 2880
DB 2821 acatggtgcttttcttcacatagagcaaaactgctttttgagaccgttagaacctctt 2880
QY 2881 agctttgtggttccctgctcaatttttatacttcttaagcaaaagtccttaggtagctt 2940
DB 2881 agctttgtggttccctgctcaatttttatacttcttaagcaaaagtccttaggtagctt 2940
QY 2941 gggatgagatgtgtgtgaaagatgtacaagagaaaaacgggaagagaggaatgaggtg 3000
DB 2941 gggatgagatgtgtgtgaaagatgtacaagagaaaaacgggaagagaggaatgaggtg 3000
QY 3001 ggggtggagaaacccatgggacagattcccatcttttagcctaactgcttcgtcattgctt 3060
DB 3001 ggggtggagaaacccatgggacagattcccatcttttagcctaactgcttcgtcattgctt 3060
QY 3061 cgtcacatcaatgcaaaaggtccctgattttgttcagcaaaacacagtcgaatgtctca 3120
DB 3061 cgtcacatcaatgcaaaaggtccctgattttgttcagcaaaacacagtcgaatgtctca 3120
QY 3121 gagtgaactttcgaaataaaattggcccaagacttttaactcgggtctttaaataatgcccac 3180
DB 3121 gagtgaactttcgaaataaaattggcccaagacttttaactcgggtctttaaataatgcccac 3180
QY 3181 attttactttgtttttcttttaataagctgggcccacatgttggaataaagctagtaatg 3240

CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.
XX
SQ Sequence 4286 BP; 1327 A; 829 C; 816 G; 1314 T; 0 other;

Query Match 99.6%; Score 4284.4; DB 21; Length 4286;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gagacattccggtggggactctggccagcccgagcaactggatcctgagacactccc 60
DB 1 gagacattccggtggggactctggccagcccgagcaactggatcctgagacactccc 60
QY 61 aggtaggatttcccgcgtgggagccttggcagagcagtgtgtggcagcccccggtg 120
DB 61 aggtaggatttcccgcgtgggagccttggcagagcagtgtgtggcagcccccggtg 120
QY 121 aggtacaaacagtgctgaacactgggaagggaactgggtacttggagcttggacattcga 180
DB 121 aggtacaaacagtgctgaacactgggaagggaactgggtacttggagcttggacattcga 180
QY 181 aacttggtctgaactgcgagcggccaccgagccttctggagcagtagcagcatg 240
DB 181 aacttggtctgaactgcgagcggccaccgagccttctggagcagtagcagcatg 240
QY 241 cagccgctccaaagtctgtgcgagcgcctgtgtgcgtgtgttcttgcctcgccctg 300
DB 241 cagccgctccaaagtctgtgcgagcgcctgtgtgcgtgtgttcttgcctcgccctg 300
QY 301 tcgcgatatcgggagagagagagccttccgccttgacaggccactccgcttttgcaa 360
DB 301 tcgcgatatcgggagagagagagccttccgccttgacaggccactccgcttttgcaa 360
QY 361 accgcagagataatgagccaccactaaagacctatggcccaagggttccacgcagct 420
DB 361 accgcagagataatgagccaccactaaagacctatggcccaagggttccacgcagct 420
QY 421 ctggcgctgctgttgccactcggaggtgcctaaaggagacaggcggcagatctccg 480
DB 421 ctggcgctgctgttgccactcggaggtgcctaaaggagacaggcggcagatctccg 480
QY 481 ccacgacacatctccctccctcccggtcccaaggaccatcgagatcagaggaattccaaa 540
DB 481 ccacgacacatctccctccctcccggtcccaaggaccatcgagatcagaggaattccaaa 540
QY 541 tacatcaaacggtgtgtctcctcctgtgtctgtcgtgggagatcgcgggaactccaca 600
DB 541 tacatcaaacggtgtgtctcctcctgtgtctgtcgtgggagatcgcgggaactccaca 600
QY 601 ctcttgagaattatctacagaacaagtgcattgcgaaacgggtcccaatatcttgatgcc 660
DB 601 ctcttgagaattatctacagaacaagtgcattgcgaaacgggtcccaatatcttgatgcc 660
QY 661 agcttgctcgtggagacctgtgcacatcgttcattgacatccctcaatgtctcacaag 720
DB 661 agcttgctcgtggagacctgtgcacatcgttcattgacatccctcaatgtctcacaag 720
QY 721 ctgctggcagagactggccatttggagctgagatgtgtaagctgtgcttcttcacacag 780
DB 721 ctgctggcagagactggccatttggagctgagatgtgtaagctgtgcttcttcacacag 780
QY 781 aaagcctccggtgggaatcactgtgtgagctctatgtgtcctgagtagttgacagatcga 840
DB 781 aaagcctccggtgggaatcactgtgtgagctctatgtgtcctgagtagttgacagatcga 840
QY 841 gctgtgtctcttggagtagaattaaaggaattgggggttcccaaaatggacagcagtagaa 900
DB 841 gctgtgtctcttggagtagaattaaaggaattgggggttcccaaaatggacagcagtagaa 900
QY 901 attgttttatttgggtgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
DB 901 attgttttatttgggtgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960

DB 901 attgttttatttgggtgtctcgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
QY 961 ataattacgatggactacaaaaggaagtattctgcgaactcgttgccttccatccggttcag 1020
DB 961 ataattacgatggactacaaaaggaagtattctgcgaactcgttgccttccatccggttcag 1020
QY 1021 aagacagcttctcagtgagttttacagacagcaaaagattgggtgcttgcagttcttat 1080
DB 1021 aagacagcttctcagtgagttttacagacagcaaaagattgggtgcttgcagttcttat 1080
QY 1081 tctcgttgcattggccatcactgcatctttttatcacataatgacctgtgaaatgttg 1140
DB 1081 tctcgttgcattggccatcactgcatctttttatcacataatgacctgtgaaatgttg 1140
QY 1141 agaaaagaaagtggcatgcagattgttttaaatgatcacctaaagcagacaggggaagt 1200
DB 1141 agaaaagaaagtggcatgcagattgttttaaatgatcacctaaagcagacaggggaagt 1200
QY 1201 gccaaaaccgtcttcttgcctggtccctgtcttcttgcctctgtcgttgccttccactc 1260
DB 1201 gccaaaaccgtcttcttgcctggtccctgtcttcttgcctctgtcgttgccttccactc 1260
QY 1261 agcagattctgaagctcactctttataatcagaatgatcccaatagatgtaacttttg 1320
DB 1261 agcagattctgaagctcactctttataatcagaatgatcccaatagatgtaacttttg 1320
QY 1321 agcttctgtgtgattggactatattgggtatcaacatggcttcaactgaattcctgcatt 1380
DB 1321 agcttctgtgtgattggactatattgggtatcaacatggcttcaactgaattcctgcatt 1380
QY 1381 aacccaattgctcgtatttgggtgagcaaaagattccaaaactgctttaaagtcatgctta 1440
DB 1381 aacccaattgctcgtatttgggtgagcaaaagattccaaaactgctttaaagtcatgctta 1440
QY 1441 tgcgtcgtggtgcagtcatttgaagaaaaacagtccttgggagaaaaagcagtcgtgctta 1500
DB 1441 tgcgtcgtggtgcagtcatttgaagaaaaacagtccttgggagaaaaagcagtcgtgctta 1500
QY 1501 aagttcaaaagcatttgcgaacacaaacaaacaaactatgtattttgcacagcacactat 1560
DB 1501 aagttcaaaagcatttgcgaacacaaacaaacaaactatgtattttgcacagcacactat 1560
QY 1561 tcttgaagaagaactatcacttgcatttcttcttcttcttcttcttcttcttcttcttctt 1620
DB 1561 tcttgaagaagaactatcacttgcatttcttcttcttcttcttcttcttcttcttcttctt 1620
QY 1621 aacaaaatgaacatttgcgaacacaaacaaacaaacaaactatgtattttgcacagcacactat 1680
DB 1621 aacaaaatgaacatttgcgaacacaaacaaacaaacaaactatgtattttgcacagcacactat 1680
QY 1681 taataattaaagtgttaattattttaacactcagctacatgacattttatgagctgt 1740
DB 1681 taataattaaagtgttaattattttaacactcagctacatgacattttatgagctgt 1740
QY 1741 ttaacgcatggaagaaataatcagtggaattaaagaaagcctcgtcgtggaagcacttaat 1800
DB 1741 ttaacgcatggaagaaataatcagtggaattaaagaaagcctcgtcgtggaagcacttaat 1800
QY 1801 tttttcagctagcacttcaacatagctcttcaacacttccaggtatttcacacacact 1860
DB 1801 tttttcagctagcacttcaacatagctcttcaacacttccaggtatttcacacacact 1860
QY 1861 taggcttaaaatgagctcactcagaaatttctattcttcttcttcttcttcttcttcttctt 1920
DB 1861 taggcttaaaatgagctcactcagaaatttctattcttcttcttcttcttcttcttcttctt 1920
QY 1921 aatcaatggactcctgatataaaggagaataaagtcaactgtcaaaacagaacttttaagt 1980
DB 1921 aatcaatggactcctgatataaaggagaataaagtcaactgtcaaaacagaacttttaagt 1980
QY 1981 aagcttaaatctactcaatttaaaattttaaattccctttaaacaacttttcaattaat 2040
DB 1981 aagcttaaatctactcaatttaaaattttaaattccctttaaacaacttttcaattaat 2040

Qy 2041 tatcacactattatcacagattgtaattagatgcaaatgagagcagtttagttgttgcatt 2100
Db 2041 tatcacactattatcacagattgtaattagatgcaaatgagagcagtttagttgttgcatt 2100
Qy 2101 ttttcggacactgaaacatttaaatgatcaggaggagtgtaacagaaagacaaagcctgt 2160
Db 2101 ttttcggacactgaaacatttaaatgatcaggaggagtgtaacagaaagacaaagcctgt 2160
Qy 2161 ttttgaaataattacacatttccactagagaagcccaaacctcagcattctgcataatgtaac 2220
Db 2161 ttttgaaataattacacatttccactagagaagcccaaacctcagcattctgcataatgtaac 2220
Qy 2221 caacatgtcacaacaacagcagcatgtacacagactggcaccatgtgccagctgaatttaaaa 2280
Db 2221 caacatgtcacaacaacagcagcatgtacacagactggcaccatgtgccagctgaatttaaaa 2280
Qy 2281 tataacttttaaaagaaaattattacatccctttacatccctcagtttaagatcaaaacctca 2340
Db 2281 tataacttttaaaagaaaattattacatccctttacatccctcagtttaagatcaaaacctca 2340
Qy 2341 caaagagaaaatagaaatgtttgaaaggctatcccaaaaagacttttttgaaatctgtcaatca 2400
Db 2341 caaagagaaaatagaaatgtttgaaaggctatcccaaaaagacttttttgaaatctgtcaatca 2400
Qy 2401 cataccctgtgaagacaatactatactacaatttttccaggatttataaaaatctcttttt 2460
Db 2401 cataccctgtgaagacaatactatactacaatttttccaggatttataaaaatctcttttt 2460
Qy 2461 tcaactatcgttagcttaaaactgtttgtttgttgcattctgtaaaacttaccatcacataca 2520
Db 2461 tcaactatcgttagcttaaaactgtttgtttgttgcattctgtaaaacttaccatcacataca 2520
Qy 2521 ctgcattgtagatgattaaatgaggcaggccctgtgctcatagctttacagatggagagat 2580
Db 2521 ctgcattgtagatgattaaatgaggcaggccctgtgctcatagctttacagatggagagat 2580
Qy 2581 gccagtgacctcataataaagactgtgaactgcctgtgagtgccacatgacaaagg 2640
Db 2581 gccagtgacctcataataaagactgtgaactgcctgtgagtgccacatgacaaagg 2640
Qy 2641 gcaggtagcaccctctcaccactgcctgtgtgttaaaatgtttctagcatatgtataat 2700
Db 2641 gcaggtagcaccctctcaccactgcctgtgtgttaaaatgtttctagcatatgtataat 2700
Qy 2701 gctatagttaaaaactactttttcaaaatcatacagatagtagacattttaaagctcacctg 2760
Db 2701 gctatagttaaaaactactttttcaaaatcatacagatagtagacattttaaagctcacctg 2760
Qy 2761 taagacttattactaaattttgtattttttgttaaatagccaatagaaaagtgtgcttg 2820
Db 2761 taagacttattactaaattttgtattttttgttaaatagccaatagaaaagtgtgcttg 2820
Qy 2821 acatgtgtcttttttccatcagaggcaaaactgctttttgagaccgtgaagaacctctt 2880
Db 2821 acatgtgtcttttttccatcagaggcaaaactgctttttgagaccgtgaagaacctctt 2880
Qy 2881 agctttgtcgttccctgcctaatttttatactctctaagcaaaagtgcotttaggtagatt 2940
Db 2881 agctttgtcgttccctgcctaatttttatactctctaagcaaaagtgcotttaggtagatt 2940
Qy 2941 gggatgagatgtgtgaaagtattgtacaagagaaaaacggagagagagaaatgaggtg 3000
Db 2941 gggatgagatgtgtgaaagtattgtacaagagaaaaacggagagagagaaatgaggtg 3000
Qy 3001 ggggtgaggaaaacccatgggacagattcccatctttagcctaagcttcgctcattgctct 3060
Db 3001 ggggtgaggaaaacccatgggacagattcccatctttagcctaagcttcgctcattgctct 3060
Qy 3061 cgtcacatcaatgcaaaaggctcctgattttgttccagcaaaacacagtcgaatgttctca 3120
Db 3061 cgtcacatcaatgcaaaaggctcctgattttgttccagcaaaacacagtcgaatgttctca 3120

Qy 3121 gagtactttcgaaataaaattgggcccaagagctttaactcggttcttaaaatagcccaa 3180
Db 3121 gagtactttcgaaataaaattgggcccaagagctttaactcggttcttaaaatagcccaa 3180
Qy 3181 attttactttgttttttcttaattaggtggccacatgttgggaaataagctagtaagt 3240
Db 3181 attttactttgttttttcttaattaggtggccacatgttgggaaataagctagtaagt 3240
Qy 3241 ttgttttctgtccaattatgaatgtgatgtgacagtaaaacccaaacccaaatgtggcca 3300
Db 3241 ttgttttctgtccaattatgaatgtgatgtgacagtaaaacccaaacccaaatgtggcca 3300
Qy 3301 gaaagaaagagcaataataataattcaattcacacacacataggtctctattataaacaccc 3360
Db 3301 gaaagaaagagcaataataataattcaattcacacacacataggtctctattataaacaccc 3360
Qy 3361 acaaaactgttttcttaatttccaaatcactttttcagaggcctgttatcatagaagt 3420
Db 3361 acaaaactgttttcttaatttccaaatcactttttcagaggcctgttatcatagaagt 3420
Qy 3421 cattttagactctcaatttttaatttaatttgaatcactcaataatttttccagtttattaa 3480
Db 3421 cattttagactctcaatttttaatttgaatcactcaataatttttccagtttattaa 3480
Qy 3481 tatatttaatttctatttaaaatttttagatttttttaccatgtcactgaatttttaca 3540
Db 3481 tatatttaatttctatttaaaatttttagatttttttaccatgtcactgaatttttaca 3540
Qy 3541 tctgtataccctttctctctccatgtcagtatcatgttctcttaatttatcttgcacaattt 3600
Db 3541 tctgtataccctttctctctccatgtcagtatcatgttctcttaatttatcttgcacaattt 3600
Qy 3601 tgaactcac 3660
Db 3601 tgaactcac 3660
Qy 3661 ttttaaaaaaaatgtttgtatcacaacttttaacacacacacacacacacacacacacacac 3720
Db 3661 ttttaaaaaaaatgtttgtatcacaacttttaacacacacacacacacacacacacacacac 3720
Qy 3721 tttcttaccatcactcaaaaac 3780
Db 3721 tttcttaccatcactcaaaaac 3780
Qy 3781 tttcttagtattaaagac 3840
Db 3781 tttcttagtattaaagac 3840
Qy 3841 cagctcaaaaagatttataaaaagatttttaacacacacacacacacacacacacacacacac 3900
Db 3841 cagctcaaaaagatttataaaaagatttttaacacacacacacacacacacacacacacacac 3900
Qy 3901 gtggatgtatgttcaaac 3960
Db 3901 gtggatgtatgttcaaac 3960
Qy 3961 gtttatacaaaaac 4020
Db 3961 gtttatacaaaaac 4020
Qy 4021 aaaaaattatatactctcggaggattttttgttgcctaaagtgtgcttagtactgattt 4080
Db 4021 aaaaaattatatactctcggaggattttttgttgcctaaagtgtgcttagtactgattt 4080
Qy 4081 tttattatgtgaagcaaaaac 4140
Db 4081 tttattatgtgaagcaaaaac 4140
Qy 4141 actgtacagac 4200
Db 4141 actgtacagac 4200
Qy 4201 caagtgagacattattttatgttaaaatatacaaatattcaagcaagtagtgaatttattcaatt 4260

Db 4201 caagtgcacattattgttaataacatacaattatcaagcaagatgaagtattcaatt 4260
Qy 4261 aaatgcccacatttctgtctctggg 4286
Db 4261 aaatgcccacatttctgtctctggg 4286

RESULT 3

AAA35162
ID AAA35162 standard; DNA; 4286 BP.

AC AAA35162..

XX 28-JUL-2000 (first entry)

XX Human adenosine receptor related polynucleotide 2nd SEQ ID NO:36.

XX Human; adenosine receptor; low adenosine antisense oligonucleotide;
KW phosphorothioate; impaired respiration; inflammation; allergy;
KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
KW lung disease; ischemic condition; pulmonary vasoconstriction; asthma;
KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX Homo sapiens.

XX WO200009525-A2.

XX 24-FEB-2000.

XX 03-AUG-1999; 99WO-US17712.

XX 03-AUG-1998; 98US-0095212.

XX (UYEC-) UNIV EAST CAROLINA.

XX Nyce JW;

XX WPI; 2000-205971/18.

XX New antisense oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers

XX Disclosure; Page 1191-1192; 1343pp; English.

XX The present invention describes a new composition comprising an
CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
CC inflammation. The ON can have antiinflammatory, antiallergic,
CC antiasthmatic, cytostatic and analgesic activities. The compositions are
CC useful for the treatment of diseases associated with inflammation,
CC impaired airways, including lung disease and diseases whose secondary
CC effects afflict the lungs of a subject. They can be used for treating
CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,
CC carcinomas, and cancers which may metastasize to the lungs, including
CC breast and prostate cancer. The reduction of the adenosine content of
CC the ONs reduces side effects. The A-containing ONs break down with the
CC release of deoxyadenosine which activates adenosine receptors causing
CC bronchoconstriction and inflammation. AAA32313 to AAA3512 represent the
CC nucleotide sequences given in the sequence listing from the present
CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
CC differ from the previously named sequences. SEQ ID NO:11 to 1880
CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
CC invention. N.B. Sequences given in the disclosure of the present

CC invention do not match up with their corresponding SEQ ID NO: sequences
CC given in the sequence listing.

XX
SQ Sequence 4286 BP; 1327 A; 829 C; 816 G; 1314 T; 0 other;

Query Match 99.6%; Score 4284.4; DB 21; Length 4286;
Best Local Similarity 100.0%; Pred. NO. 0;
Matches 4285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 gagacattccggtgggactctgcccagcccgagcaactggagctctgagacactccc 60
Db 1 gagacattccggtgggactctgcccagcccgagcaactggagctctgagacactccc 60

Qy 61 aggtaggcatttgcgccggtgggacgacctgcccagagcagtgctggcagccccccgtgg 120
Db 61 aggtaggcatttgcgccggtgggacgacctgcccagagcagtgctggcagccccccgtgg 120

Qy 121 aggtatcacacagtggtgaaactgggaaggaaactggtacttgagctctggacatctga 180
Db 121 aggtatcacacagtggtgaaactgggaaggaaactggtacttgagctctggacatctga 180

Qy 181 aacttgctctgaaactgcgagcgccacccgagcctctctggagcaggtagcagcatg 240
Db 181 aacttgctctgaaactgcgagcgccacccgagcctctctggagcaggtagcagcatg 240

Qy 241 cagcgcctccaaagtctgtgcgagcgccctggtgctgctggttcttgcctgcggcctg 300
Db 241 cagcgcctccaaagtctgtgcgagcgccctggtgctgctggttcttgcctgcggcctg 300

Qy 301 tcgcggtatctgggagagagagaggttcccgctctgacaggggccaactccgcttttgc 360
Db 301 tcgcggtatctgggagagagagaggttcccgctctgacaggggccaactccgcttttgc 360

Qy 361 accgacagataatgacgccaccacctaagaccttatggcccaggggttccaaagccag 420
Db 361 accgacagataatgacgccaccacctaagaccttatggcccaggggttccaaagccag 420

Qy 421 ctggcgcggtctgttggcaccctgcgaggtgctaaagagagacagcgagcgaggtcttc 480
Db 421 ctggcgcggtctgttggcaccctgcgaggtgctaaagagagacagcgaggtcttc 480

Qy 481 ccaagcaccatctccctcccccgtgccaaggaccatcgagatcaaggagacttcaaa 540
Db 481 ccaagcaccatctccctcccccgtgccaaggaccatcgagatcaaggagacttcaaa 540

Qy 541 tacatcaacacggtgtgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600
Db 541 tacatcaacacggtgtgtctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 600

Qy 601 cttctgagaattatctcaagaacaagtgcagtcgaaacggtcccaatacttctgacgcc 660
Db 601 cttctgagaattatctcaagaacaagtgcagtcgaaacggtcccaatacttctgacgcc 660

Qy 661 agcttgctctgggagacctgctgcacatcgctcatattgacatccctataatgtctaca 720
Db 661 agcttgctctgggagacctgctgcacatcgctcatattgacatccctataatgtctaca 720

Qy 721 ctgctggcagagactggccatttggagctgagatgtgtaagctgggtgcttcttcacag 780
Db 721 ctgctggcagagactggccatttggagctgagatgtgtaagctgggtgcttcttcacag 780

Qy 781 aaagcctcccggtgggaatcactgtgctgagctctatgtctctgagtatggacagatctga 840
Db 781 aaagcctcccggtgggaatcactgtgctgagctctatgtctctgagtatggacagatctga 840

Qy 841 gctgtgtcttcttggagtagaataaaaggaaatgggggttccaaatggacagcagtagaa 900
Db 841 gctgtgtcttcttggagtagaataaaaggaaatgggggttccaaatggacagcagtagaa 900

Qy 901 attgtttgtattgggtggtgctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960
Db 901 attgtttgtattgggtggtgctctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 960

QY	961	ataattacagatgactacaaaagaagtattctgcgaatctgtgcttcatccggttcag	1020	QY	2041	tatcaactattatcatcagattgtaattagatgcaaatgagagagcagtttagttgttcac	2100
Db	961	ataattacagatgactacaaaagaagtattctgcgaatctgtgcttcatccggttcag	1020	Db	2041	tatcaactattatcatcagattgtaattagatgcaaatgagagagcagtttagttgttcac	2100
QY	1021	aagacagcttccatgcagttttacaagacagcaaaagattggtgcgttcagtttctat	1080	QY	2101	ttttcgacactggaacattttaaatgatacaggaggagtaacagaaagagcaggtcgt	2160
Db	1021	aagacagcttccatgcagttttacaagacagcaaaagattggtgcgttcagtttctat	1080	Db	2101	ttttcgacactggaacattttaaatgatacaggaggagtaacagaaagagcaggtcgt	2160
QY	1081	ttctgcttgccattggccatacactcattttttttatacactaatgacctgtgaaatgttg	1140	QY	2161	ttttgaaatacattacacitttactagaagcccaaaacctcagcattctgcaatatgtaac	2220
Db	1081	ttctgcttgccattggccatacactcattttttttatacactaatgacctgtgaaatgttg	1140	Db	2161	ttttgaaatacattacacitttactagaagcccaaaacctcagcattctgcaatatgtaac	2220
QY	1141	agaagaaaaagtgccatgcagattgcttttaaatgatcacctaaagcagagacgggaagtg	1200	QY	2221	caacatgtaacaaacagcagcatgttaacagactggcacatgtgccagctggaatttaaaa	2280
Db	1141	agaagaaaaagtgccatgcagattgcttttaaatgatcacctaaagcagagacgggaagtg	1200	Db	2221	caacatgtaacaaacagcagcatgttaacagactggcacatgtgccagctggaatttaaaa	2280
QY	1201	gccaaaaacgctgttttgcctgtcttcttgcctctgtgcttcccttccacctc	1260	QY	2281	tataatacttttaaaagaaaaattattacatcctttacatttcagtttaagatcaaacctca	2340
Db	1201	gccaaaaacgctgttttgcctgtcttcttgcctctgtgcttcccttccacctc	1260	Db	2281	tataatacttttaaaagaaaaattattacatcctttacatttcagtttaagatcaaacctca	2340
QY	1261	agcagattctgaagctcactctttataatcagaatgatcccaatagatgtgaactttg	1320	QY	2341	caaaagaaaatagaaatgtttgaaaggctatcccaaaagacttttttgaaactgtcatca	2400
Db	1261	agcagattctgaagctcactctttataatcagaatgatcccaatagatgtgaactttg	1320	Db	2341	caaaagaaaatagaaatgtttgaaaggctatcccaaaagacttttttgaaactgtcatca	2400
QY	1321	agctttctgtgttattggaactatattggtatacaacatggcttccactgaattccctgcatt	1380	QY	2401	cataccctgtgaagacaatactatctacaaattttttcaggattattaaaaactctctttt	2460
Db	1321	agctttctgtgttattggaactatattggtatacaacatggcttccactgaattccctgcatt	1380	Db	2401	cataccctgtgaagacaatactatctacaaattttttcaggattattaaaaactctctttt	2460
QY	1381	aacccaaattgctctgatttggtagcaaaagattcaaaaactgttttaagtcagcttca	1440	QY	2461	tcactatcgtagcttaaaactctgtttggtttgtctatctctgtaaaactactaccatataca	2520
Db	1381	aacccaaattgctctgatttggtagcaaaagattcaaaaactgttttaagtcagcttca	1440	Db	2461	tcactatcgtagcttaaaactctgtttggtttgtctatctctgtaaaactactaccatataca	2520
QY	1441	tgctgctggtgccagtcatttgaaagaaaaacagtccttgagagaaaaacgagtcgctgctta	1500	QY	2521	ctgcatgtagatgattaaatgaggcgagccctgtgctcattagcttttacgatatgagagat	2580
Db	1441	tgctgctggtgccagtcatttgaaagaaaaacagtccttgagagaaaaacgagtcgctgctta	1500	Db	2521	ctgcatgtagatgattaaatgaggcgagccctgtgctcattagcttttacgatatgagagat	2580
QY	1501	aagttcaaaagctaaatgatacagcagatgacaaactccgttccagtaataaaacacagctca	1560	QY	2581	gccagtgacctcataataaagactgtgaactgctggtgagtcgagtcgcacatgacaaagg	2640
Db	1501	aagttcaaaagctaaatgatacagcagatgacaaactccgttccagtaataaaacacagctca	1560	Db	2581	gccagtgacctcataataaagactgtgaactgctggtgagtcgagtcgcacatgacaaagg	2640
QY	1561	tcttgaaagaagaactatttcaactgtatttctatttcttttataattggaccgaagtcattaa	1620	QY	2641	gcaggtagcacctctctcacccatgctgtggtttaaattggtttctagcatagtataat	2700
Db	1561	tcttgaaagaagaactatttcaactgtatttctatttcttttataattggaccgaagtcattaa	1620	Db	2641	gcaggtagcacctctctcacccatgctgtggtttaaattggtttctagcatagtataat	2700
QY	1621	aacaaaaatgaacatttgcacaaacaaacaaaaaactatgtatttgacagcacactat	1680	QY	2701	gctatagtttaaaataactatttttcaaaaaatacacagattagtaacatttaacagctacctg	2760
Db	1621	aacaaaaatgaacatttgcacaaacaaacaaaaaactatgtatttgacagcacactat	1680	Db	2701	gctatagtttaaaataactatttttcaaaaaatacacagattagtaacatttaacagctacctg	2760
QY	1681	taaaataataagtgtaatttttaacactcacagctacatatgacatttttatgacctgt	1740	QY	2761	taaaagcttattactaattttgtattatttttgaataagccaatagaaagtttgccttg	2820
Db	1681	taaaataataagtgtaatttttaacactcacagctacatatgacatttttatgacctgt	1740	Db	2761	taaaagcttattactaattttgtattatttttgaataagccaatagaaagtttgccttg	2820
QY	1741	ttacggcatgaaagaaaatcagtggggaattaaagaagcctgcgtgaaagcacttaat	1800	QY	2821	acatgggtgctttcttctcatcagaggcaaaactgctttttgagaccgtgaagaacctctt	2880
Db	1741	ttacggcatgaaagaaaatcagtggggaattaaagaagcctgcgtgaaagcacttaat	1800	Db	2821	acatgggtgctttcttctcatcagaggcaaaactgctttttgagaccgtgaagaacctctt	2880
QY	1801	tttttacagtttagcaactcaacatagctctttaaacaacttccaggattattcacacaact	1860	QY	2881	agctttgctgcttcctgcctaaatttttatactcttcaagcaaaagtccttagatagctt	2940
Db	1801	tttttacagtttagcaactcaacatagctctttaaacaacttccaggattattcacacaact	1860	Db	2881	agctttgctgcttcctgcctaaatttttatactcttcaagcaaaagtccttagatagctt	2940
QY	1861	tagccttaaaatgagctcactcagaattttctattctttcttaaaagagatttttta	1920	QY	2941	gggatgagatgctgtgaaagtattgtacaagagaaacggaagagagagaaatgaggtg	3000
Db	1861	tagccttaaaatgagctcactcagaattttctattctttcttaaaagagatttttta	1920	Db	2941	gggatgagatgctgtgaaagtattgtacaagagaaacggaagagagagaaatgaggtg	3000
QY	1921	aatcaatgggactctgataaaaggaagaataaagtcactgtaaaaacagaacttttaagt	1980	QY	3001	gggttggagaaacccatgggacagattccccattcttagcctaaagcttcgtcatgacct	3060
Db	1921	aatcaatgggactctgataaaaggaagaataaagtcactgtaaaaacagaacttttaagt	1980	Db	3001	gggttggagaaacccatgggacagattccccattcttagcctaaagcttcgtcatgacct	3060
QY	1981	aagcttaaatctactcaatttaaaatttttaaaatccttttaaaacaaacttttcaattaat	2040	QY	3061	ogtcacatacaatgcaaaaggtcctgattttgttccagcaaaacacacagtcgaattgtctca	3120
Db	1981	aagcttaaatctactcaatttaaaatttttaaaatccttttaaaacaaacttttcaattaat	2040	Db	3061	ogtcacatacaatgcaaaaggtcctgattttgttccagcaaaacacacagtcgaattgtctca	3120
				QY	3121	gagtgactttcgaaaaataattgggcccgaagactttaactcggtcttataaaatatgccaa	3180

CC allergies, asthma, impeded respiration, respiratory distress syndrome
CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
CC fragments and antisense oligonucleotides used in the exemplification of
CC the present invention.

XX Sequence 13611 BP; 3676 A; 3007 C; 3056 G; 3868 T; 4 other;

Query Match 99.6%; Score 4284.4; DB 21; Length 13611;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 4285; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 gagacattccgggtggggaactctggccagcccgagcaactggatcctcgagagcaactccc 60
|||||
DB 1873 gagacattccgggtggggaactctggccagcccgagcaactggatcctcgagagcaactccc 1932
QY 61 aggtaggcatttcccgggtgggagcccttccagagcagtgctggtggcagcccccctgg 120
|||||
DB 1933 aggtaggcatttcccgggtgggagcccttccagagcagtgctggtggcagcccccctgg 1992
QY 121 aggatcaacacagtggtgaacactgggaagaaactggtacttggagtctggacatctga 180
|||||
DB 1993 aggatcaacacagtggtgaacactgggaagaaactggtacttggagtctggacatctga 2052
QY 181 aacttggctctgaactcgagcggccacccgagccctcttgagcaggttagcagcatg 240
|||||
DB 2053 aacttggctctgaactcgagcggccacccgagccctcttgagcaggttagcagcatg 2112
QY 241 cagcgcctccaaactctgtgcgagcggccctgttgcgtgggttcttgcctgcggcctg 300
|||||
DB 2113 cagcgcctccaaactctgtgcgagcggccctgttgcgtgggttcttgcctgcggcctg 2172
QY 301 tcgcggtctgggagagagagaggttcccgctgcagggccactccgcttttgcaa 360
|||||
DB 2173 tcgcggtctgggagagagagaggttcccgctgcagggccactccgcttttgcaa 2232
QY 361 accgcagagataatgagccaccactaaagaccttatggcccaagggttccaaagccagt 420
|||||
DB 2233 accgcagagataatgagccaccactaaagaccttatggcccaagggttccaaagccagt 2292
QY 421 ctggcgcggtctgttggcaactcgagagtgctctaaagagacagagcggcaggtatctccg 480
|||||
DB 2293 ctggcgcggtctgttggcaactcgagagtgctctaaagagacagagcggcaggtatctccg 2352
QY 481 ccacgacacattccctcccccgtgcgaaggaccactcagagatacgaagagacttccaaa 540
|||||
DB 2353 ccacgacacattccctcccccgtgcgaaggaccactcagagatacgaagagacttccaaa 2412
QY 541 tacatcaacacggttctgtctgcttctgttctgtgctggggtatcctgggaactccaca 600
|||||
DB 2413 tacatcaacacggttctgtctgcttctgttctgtgctggggtatcctgggaactccaca 2472
QY 601 ctcttgagaattatctacaagaacaagtgcagtcgaaacggtcccaaatacttctgatcgcc 660
|||||
DB 2473 ctcttgagaattatctacaagaacaagtgcagtcgaaacggtcccaaatacttctgatcgcc 2532
QY 661 agcttggctctgggagacactgtgtgcacatgtcatttgacataccctataccatgtctacaag 720
|||||
DB 2533 agcttggctctgggagacactgtgtgcacatgtcatttgacataccctataccatgtctacaag 2592
QY 721 ctgctgcagagagactggccaatttggagctgagatgtgtaagctgggtcccttctcatacag 780
|||||
DB 2593 ctgctgcagagagactggccaatttggagctgagatgtgtaagctgggtcccttctcatacag 2652
QY 781 aaagccctccggtggaatacactgtgtgactgtatgtctgtgactgttgcagatatcga 840
|||||
DB 2653 aaagccctccggtggaatacactgtgtgactgtatgtctgtgactgttgcagatatcga 2712
QY 841 gctgtgtcttctgtgagtagaattaaaggaattggggttccaaaatggacagcagtagaa 900
|||||

DB 2713 gctgtgtcttctgtgagtagaattaaagggaattggggttccaaaatggacagcagtagaa 2772
QY 901 attgttttgatttgggtgtgtctctgtgttctgtgtgtccctgaagccatagttttgat 960
|||||
DB 2773 attgttttgatttgggtgtgtctctgtgttctgtgtgtccctgaagccatagttttgat 2832
QY 961 ataatcagtagtgactacaaaaggatttatctgcgaatctgtctgttcttccatcccggtcag 1020
|||||
DB 2833 ataatcagtagtgactacaaaaggatttatctgcgaatctgtctgttcttccatcccggtcag 2892
QY 1021 aagcagcgtttctcagtcagttttacaagacagcaaaagattggtggctgttctcagtttctat 1080
|||||
DB 2893 aagcagcgtttctcagtcagttttacaagacagcaaaagattggtggctgttctcagtttctat 2952
QY 1081 ttctgtctgccatttggccatcactgcattttttttatcacactaatgacctgtgaaatgttg 1140
|||||
DB 2953 ttctgtctgccatttggccatcactgcattttttttatcacactaatgacctgtgaaatgttg 3012
QY 1141 agaaagaaaagtggcatgcagattctttaaataatgatcactaaaagcagagagcgggaagtg 1200
|||||
DB 3013 agaaagaaaagtggcatgcagattctttaaataatgatcactaaaagcagagagcgggaagtg 3072
QY 1201 gccaaaaccgtcttctgtcgtgctctgtcttctgtcctctgtcgttcccttccacctc 1260
|||||
DB 3073 gccaaaaccgtcttctgtcgtgctctgtcttctgtcctctgtcgttcccttccacctc 3132
QY 1261 agcaggaattctgaagctcactcttataatcagaatgatcccaatagatgtgaaacttttg 1320
|||||
DB 3133 agcaggaattctgaagctcactcttataatcagaatgatcccaatagatgtgaaacttttg 3192
QY 1321 agcttctgttgggtatttggactatatttgggtatcaacatggcttccactgaattccctgcatt 1380
|||||
DB 3193 agcttctgttgggtatttggactatatttgggtatcaacatggcttccactgaattccctgcatt 3252
QY 1381 aaccocaaatgctctgtatttgggtgagcaaaagattcaaaaactgctttaaagtcagtctta 1440
|||||
DB 3253 aaccocaaatgctctgtatttgggtgagcaaaagattcaaaaactgctttaaagtcagtctta 3312
QY 1441 tgcgtcgtggtgcagtcatttgaagaaaacacagtccttggagaaaacagcagtcgctgctta 1500
|||||
DB 3313 tgcgtcgtggtgcagtcatttgaagaaaacacagtccttggagaaaacagcagtcgctgctta 3372
QY 1501 aagttcaaaagtaatacagcagatgcacgaatgcacaaacttccgttccagtaataaaacagctca 1560
|||||
DB 3373 aagttcaaaagtaatacagcagatgcacgaatgcacaaacttccgttccagtaataaaacagctca 3432
QY 1561 tcttgaagagaagaactatcactgtatttccatttcttatttatttggaccgagatcattaa 1620
|||||
DB 3433 tcttgaagagaagaactatcactgtatttccatttcttatttatttggaccgagatcattaa 3492
QY 1621 aacaaaatgaaacatttgcacaaaacacaaaactatgtatttgcacagcacactat 1680
|||||
DB 3493 aacaaaatgaaacatttgcacaaaacacaaaactatgtatttgcacagcacactat 3552
QY 1681 taaaataattaagtgttaatttttaacactcagctacatgacatattttatggagctgt 1740
|||||
DB 3553 taaaataattaagtgttaatttttaacactcagctacatgacatattttatggagctgt 3612
QY 1741 ttacggcagtgaaagaaaatacagtgagggaatgaagaagcctcgtcgtgaaagcacttaat 1800
|||||
DB 3613 ttacggcagtgaaagaaaatacagtgagggaatgaagaagcctcgtcgtgaaagcacttaat 3672
QY 1801 tttttacagtttagcacttcaacatagctcttaaacacttccagggtatttccacacacact 1860
|||||
DB 3673 tttttacagtttagcacttcaacatagctcttaaacacttccagggtatttccacacacact 3732
QY 1861 taggccttaaaaatgagctcactcagaattctcttcttcttcttaaaaagagattttttta 1920
|||||
DB 3733 taggccttaaaaatgagctcactcagaattctcttcttcttcttaaaaagagattttttta 3792
QY 1921 aatcaatgggactctgatataaaggaagaaataagtcactgttaaaacagaaacttttaaatg 1980
|||||
DB 3793 aatcaatgggactctgatataaaggaagaaataagtcactgttaaaacagaaacttttaaatg 3852

6013	acgtacagacactaattcattaaatctactaattgatttttaaagaataataaattgga	6072
Db		
4201	caagtggacattatttattgttaaataatacaattatcaagcaagtatgaagtattatcaatt	4260
Qy		
6073	caagtggacattatttattgttaaataatacaattatcaagcaagtatgaagtattatcaatt	6132
Db		

Qy 4261 aaatgccacatttctggtctctctggg 4286
 . |||||
 Db 6133 aaatgccacatttctggtctctctggg 6158

RESULT 5
AAA35166
ID AAA35166 standard: DNA: 13612 BP.

...	28-JUL-2000 (first entry)	Human adenosine receptor related polynucleotide 2nd SEQ ID NO:40.
DT		
XX		
DE		

Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

OS Homo sapiens.

PN WO200009525-A2.

24-FEB-2000

03-AUG-1999: 99W0-11S17712

XX
PR 03-AUG-1998 0805-0095212XX
BA (UYEC -) UNITV EAST CAROT INA

XX	No.	TJ.
DT		

XX
DE
WET. 2000 205071 110

XX New antisease oligonucleotides useful for treating e.g. pulmonary
PT vasoconstriction, inflammation, allergies, asthma, hypertension,
PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
PT cancers -

PS Disclosure: page 1194-1197: 1343no. English

The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airflow, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, leukoemias, lymphomas, pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32313 to AAA353312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last


```
Db 4933 cgtcaccaatgcaaaaggctccgtatgtttgttcagacaaacacagtgcaatgtctca 4992
Qy 3121 gagtgactttcgaaataaattgggcccagagctttaactcgttcttaaaatattgcccac 3180
Db 4993 gagtgactttcgaaataaattgggcccagagctttaactcgttcttaaaatattgcccac 5052
Qy 3181 attttactttgttttttttaattagctgggcccacatgtttggaataaagctagtaatg 3240
Db 5053 attttactttgttttttttaattagctgggcccacatgtttggaataaagctagtaatg 5112
Qy 3241 ttgtttctgtcaatattgaatgtagtggttacagtaaacaccccaacaaatgtggcca 3300
Db 5113 ttgtttctgtcaatattgaatgtagtggttacagtaaacaccccaacaaatgtggcca 5172
Qy 3301 gaaagaaagagcaataataataattcacacacacatggttctatttataaatcaccc 3360
Db 5173 gaaagaaagagcaataataataattcacacacacatggttctatttataaatcaccc 5232
Qy 3361 acaactgtttctttaatttcacccaactttttcagagggcgtgttacaatagaagt 3420
Db 5233 acaactgtttctttaatttcacccaactttttcagagggcgtgttacaatagaagt 5292
Qy 3421 cattttagactctcaattttaataattttgaatcactaatttttcacagtttattaa 3480
Db 5293 cattttagactctcaattttaataattttgaatcactaatttttcacagtttattaa 5352
Qy 3481 tataatttaattctatttaatttagattattttttattacatgactgaatttttaca 3540
Db 5353 tataatttaattctatttaatttagattattttttattacatgactgaatttttaca 5412
Qy 3541 tctgtacaccttttccctctccatgctcagtcacgtatctcttaattcttgcacaaatt 3600
Db 5413 tctgtacaccttttccctctccatgctcagtcacgtatctcttaattcttgcacaaatt 5472
Qy 3601 tgaacttacacacacaaagcaacttgcatatttattataataaaattgcatcagtggtt 3660
Db 5473 tgaacttacacacacaaagcaacttgcatatttattataataaaattgcatcagtggtt 5532
Qy 3661 tttaaaaaaattgttgattcaaaactttaacatactactgataagtaagaacaaattaat 3720
Db 5533 tttaaaaaaattgttgattcaaaactttaacatactactgataagtaagaacaaattaat 5592
Qy 3721 tctttacatactcaaaacacacagtagaagagtgctatogttcaacttcaaaacatgt 3780
Db 5593 tctttacatactcaaaacacagtagaagagtgctatogttcaacttcaaaacatgt 5652
Qy 3781 tctctagtataaggaactttaatatagacaacagacaaaattattgttaacatggatgta 3840
Db 5653 tctctagtataaggaactttaatatagacaacagacaaaattattgttaacatggatgta 5712
Qy 3841 cagctcaaaagattataaaagattttaaacatttttcccttatttccactgctaatt 3900
Db 5713 cagctcaaaagattataaaagattttaaacatttttcccttatttccactgctaatt 5772
Qy 3901 gtggatgtatgttcaaacacaccttttagtattgtatagcttacaatgcccagaagaaataca 3960
Db 5773 gtggatgtatgttcaaacacaccttttagtattgtatagcttacaatgcccagaagaaataca 5832
Qy 3961 gtttatagcaaaacatgggtatgctgtagctaaactttataaaagtgtataataacaatgt 4020
Db 5833 gtttatagcaaaacatgggtatgctgtagctaaactttataaaagtgtataataacaatgt 5892
Qy 4021 aaaaaattatatctgtggagattttttgttgcttaaaagtgtactagttactgattt 4080
Db 5893 aaaaaattatatctgtggagattttttgttgcttaaaagtgtactagttactgattt 5952
Qy 4081 ttattatgtaagcaaaccaataaaaaatttaagtttttttaaacactacattattttc 4140
Db 5953 ttattatgtaagcaaaccaataaaaaatttaagtttttttaaacactacattattttc 6012
Qy 4141 actgtacagacactaattcattataactaattgattgttttaaaagaaataataatgtga 4200
```

```
Db 6013 actgtacagacactaattcaataactaattgattgttttaaaagaaataataatgtga 6072
Qy 4201 caagtggaactattattgtttaaataatacaattatacaagcaagtatgaagttattcaatt 4260
Db 6073 caagtggaactattattgtttaaataatacaattatacaagcaagtaagaagttattcaatt 6132
Qy 4261 aaaaatggcacattttctgtctctggg 4286
Db 6133 aaaaatggcacattttctgtctctggg 6158

RESULT 6
AAV17875
ID AAV17875 standard; cDNA to mRNA; 1873 BP.
XX
AC AAV17875;
XX
DT 03-AUG-1998 (first entry)
XX
DE Homo sapiens endothelin B receptor gene.
XX
KW dystroglycan; hDG; progressive muscular dystrophy;
KW endothelin B receptor; alpha-sarcoglycan; Hirschsprung's disease; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 44..1204
FT /tag= a
FT /product= dystroglycan
XX
PN EP825265-A1.
XX
PD 25-FEB-1998.
XX
PF 14-AUG-1997; 97EP-0114019.
XX
PR 16-AUG-1996; 96JP-0216506.
XX
PA (RIKA ) INST PHYSICAL & CHEM RES.
PA (RIKA ) RIKAGAKU KENKYUSHO.
XX
PI Hanaoka F, Sakamoto A;
XX
DR WPI; 1998-132258/13.
XX
PT cDNA reverse transcription and amplification - using different
PT reverse transcription and 3' amplification primers
XX
PS Example 3; Page 19-21; 25pp; English.
XX
CC The sequence is that encoding human endothelin B receptor which is
CC implicated in Hirschsprung's disease.
XX
SQ Sequence 1873 BP; 490 A; 434 C; 438 G; 511 T; 0 other;
```

Query Match 39.3%; Score 1691.8; DB 19; Length 1873;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1696; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

```
Qy 178 tgaacttggtctgaaactcggagcgccacccggagccctctctggagcaggtagcagc 237
Db 171 tgtctctaggtctgaaactcggagcgccacccggagccctctctggagcaggtagcagc 230
Qy 238 atgcagcgcgcctcaagtctgtcggaacgcgccttggtcgctggttcttccctgcgcgc 297
Db 231 atgcagcgcgcctcaagtctgtcggaacgcgccttggtcgctggttcttccctgcgcgc 290
Qy 298 ctgtcgcggtctctgggagagagagaggttcccgctgacagggccactccgcttttg 357
Db 291 ctgtcgcggtctctgggagagagagaggttcccgctgacagggccactccgcttttg 350
```


Db	1731	tgtttacgcatggaagaatacagtg9gaataaagaagcctcgtcgtgaaagcactt	1790
Qy	1798	aatttttacagttacacattcaacatagctcttaacaacttcacagattatcacacaac	1857
Db	1791	aatttttacagttacacattcaacatagctcttaacaacttcacagattatcacacaac	1850
Qy	1858	acttaggcttaaaaatgagctc	1879
Db	1851	acttaggcttaaaaatgagctc	1872
RESULT	8		
AAA35161			
ID	AAA35161	standard; DNA; 1872 BP.	
XX	XX	AC	XX
XX	XX	AAA35161;	
DT	28-JUL-2000	(first entry)	
XX	XX		
DE	Human	adenosine receptor related polynucleotide 2nd SEQ ID NO:35.	
XX	XX		
KW	Human; adenosine receptor; low adenosine antisense oligonucleotide; phosphorothioate; impaired respiration; inflammation; allergy; allergic disease; bronchoconstriction; inhibitor; antiinflammatory; antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway; lung disease; ischaemic condition; pulmonary vasoconstriction; asthma; respiratory distress syndrome; pain; cystic fibrosis; emphysema; pulmonary hypertension; chronic obstructive pulmonary disease; COPD; cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.		
XX	XX		
OS	Homo sapiens.		
XX	XX		
PN	WO200009525-A2.		
XX	XX		
PD	24-FEB-2000.		
XX	XX		
PF	03-AUG-1999; 99WO-US17712.		
XX	XX		
PR	03-AUG-1998; 98US-0095212.		
XX	XX		
PA	(UYEC-) UNIV EAST CAROLINA.		
XX	XX		
PI	Nyce JW;		
XX	XX		
DR	WPI; 2000-205971/18.		
XX	XX		
PT	New antisense oligonucleotides useful for treating e.g. pulmonary vasoconstriction, inflammation, allergies, asthma, hypertension, bronchitis, emphysema, respiratory distress syndrome, ischemia or cancers		
PT			
XX	XX		
PS	Disclosure; Page 1190-1191; 1343pp; English.		
XX	XX		
CC	The present invention describes a new composition comprising an antisense oligonucleotide (ON) with low adenosine (up to 15%), which targets nucleic acids involved in bronchoconstriction, allergies, and/or inflammation. The ON can have antiinflammatory, antiallergic, antiasthmatic, cytostatic and analgesic activities. The compositions are useful for the treatment of diseases associated with inflammation, impaired airways, including lung disease and diseases whose secondary effects afflict the lungs of a subject. They can be used for treating e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma, impaired respiration, respiratory distress syndrome, pain, cystic fibrosis, pulmonary hypertension, emphysema, chronic obstructive pulmonary disease (COPD), and cancers such as leukaemias, lymphomas, carcinomas, and cancers which may metastasise to the lungs, including breast and prostate cancer. The reduction of the adenosine content of the ONs reduces side effects. The A-containing ONs break down with the release of deoxyadenosine which activates adenosine receptors causing bronchoconstriction and inflammation. AAA32133 to AAA35312 represent the nucleotide sequences given in the sequence listing from the present invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185 sequences are also called SEQ ID NO:1 to 185, but the sequences		

CC	differ from the previously named sequences. SEQ ID NO:11 to 1680
CC	(AAA32323 to AAA33992) are specifically claimed ONs from the present
CC	invention. N.B. Sequences given in the disclosure of the present
CC	invention do not match up with their corresponding SEQ ID NO: sequences
CC	given in the sequence listing.
XX	
SO	Sequence 1872 BP; 489 A; 434 C; 438 G; 511 T; 0 other;
	Query Match 39.3%; Score 1690.8; DB 21; Length 1872;
	Best Local Similarity 99.68; Pred. No. 0;
	Matches 1695; Conservative 0; Mismatches 7; Indels 0; Gaps
Qy	178 tgaacttggctctgaactcggagcggccacgcggacgctctctggacgaggtagcagc 237
Db	171 tgtctaggctctgaactcggagcggccacgcggacgctctctggacgaggtagcagc 230
Qy	238 atgcagcgcgcctcaagtctgtgcgacgcgcgcctgtgtgcgttgcttgcctgcgcgc 297
Db	231 atgcagcgcgcctcaagtctgtgcgacgcgcgcctgtgtgcgttgcttgcctgcgcgc 290
Qy	298 ctgtcgcggaactctggggagagagagagcttccgcctgcagcggccactccgcctttg 357
Db	291 ctgtcgcggaactctggggagagagagagcttccgcctgcagcggccactccgcctttg 350
Qy	358 caaacgcgagagataatgacgccaccaccactaagaaccttatggcccaagggttccaaagcc 417
Db	351 caaacgcgagagataatgacgccaccaccactaagaaccttatggcccaagggttccaaagcc 410
Qy	418 agtctgcgcggttcgttgcacctcggagtgctctaaaggagacagacgacgagatct 477
Db	411 agtctgcgcggttcgttgcacccgcggaggtgctctaaaggagacagacgacgagatct 470
Qy	478 ccgcacgcacactctccctcccccctgcgaaggaccctcgagatcgaaggacacttc 537
Db	471 ccgcacgcacactctccctcccccctgcgaaggaccctcgagatcgaaggacacttc 530
Qy	538 aaatcacatcaacaggttgctgcgttcgtgtgctgcgtgcgtgcgtgcgtgcgtgcgtgc 597
Db	531 aaatcacatcaacaggttgctgcgttcgtgtgctgcgtgcgtgcgtgcgtgcgtgcgtgc 590
Qy	598 acactctcgaataattctcacaagaacaagtcgatcgcaaacggtcccaatctctgac 657
Db	591 acactctcgaataattctcacaagaacaagtcgatcgcaaacggtcccaatctctgac 650
Qy	658 gccagcttggtctcgggagacctgctgcacatcgtctatttgacatccctatacattgtctac 717
Db	651 gccagcttggtctcgggagacctgctgcacatcgtctatttgacatccctatacattgtctac 710
Qy	718 aagctgctgcgagaggactgcccatttggagctcgagatggttaagctgggtgcctttcata 777
Db	711 aagctgctgcgagaggactgcccatttggagctcgagatggttaagctgggtgcctttcata 770
Qy	778 cagaagcctccgtcgggaatacactgctgagctctatgtgctcgtgagattattgcacagatat 837
Db	771 cagaagcctccgtcgggaatacactgctgagctctatgtgctcgtgagattattgcacagatat 830
Qy	838 cgagctgtgctctctggagtagaattaaaggaaattgggggttccaaaatggacagcagta 897
Db	831 cgagctgtgctctctggagtagaattaaaggaaattgggggttccaaaatggacagcagta 890
Qy	898 gaaattgttttgatttgggttgctctgtggttctgactgtccctgaagccatagggttt 957
Db	891 gaaattgttttgatttgggttgctctgtggttctgactgtccctgaagccatagggttt 950
Qy	958 gataataatcagatggactacaaggaaattctctgcgaattctgctctgtgcttcacccggt 1017
Db	951 gatataatcagatggactacaaggaaattctctgcgaattctgctctgtgcttcacccggt 1010
Qy	1018 cagaagacagctttcatcgcaggttttacaagacagcaaaaagattggttggtcgttcagtttc 1077
Db	1011 cagaagacagctttcatcgcaggttttacaagacagcaaaaagattggttggtcgttcagtttc 1070

QY 1446 ctgggtccagtcatttgaaagaaacagctcttgaggagaaagcagctgctgcttaagt 1505
|||||
Db 1496 ctgggtccagtcatttgaaagaaacagctcttgaggagaaagcagctgctgcttaagt 1555
|||||
QY 1506 caaagctaatacagatgacatgacaaactccgttccagtaataataacagctcatcttg 1565
|||||
Db 1556 caaagctaatacagatgacatgacaaactccgttccagtaataataacagctcatcttg 1615
|||||

QY 1566 aaagaagaa 1574
|||||

Db 1616 aaagaagaa 1624
|||||

RESULT 12

ABI97988
ID ABI97988 standard; cDNA; 1329 BP.

XX
AC ABI97988;

XX
DT 18-FEB-2002 (first entry)

XX
DE Non-endogenous human GPCR cDNA, SEQ ID NO: 496.

XX
KW Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
constitutively activated GPCR; agonist; disease; ss.

XX
OS Homo sapiens.

XX
OS Synthetic.

XX
PN WO200177172-A2.

XX
XX 18-OCT-2001.

XX
XX 05-APR-2001; 2001WO-US11098.

XX
XX 07-APR-2000; 2000US-195747P.

XX
PA (AREN-) ARENA PHARM INC.

XX
PI Lehmann-Bruinsma K, Liaw CW, Lin I;

DR
WPI: 2001-648759/74.

DR
P-PSDB; ABB56352.

XX
PT Identifying agonists of G protein-coupled receptors (GPCRs) for use in
disease treatment, comprises contacting candidate compounds with
versions of GPCRs -

XX
PS Example 2; Page 297; 394pp; English.

XX
CC The invention relates to G protein-coupled receptors (GPCRs) for which
the endogenous ligand has been identified. Non-endogenous
constitutively activated versions of known GPCRs are used in the
invention for the direct identification of candidate compounds as
receptor agonists, inverse agonists or partial agonists. Such
agonists are useful as therapeutic agents for diseases or disorders
associated with GPCRs. The present sequence encodes a non-endogenous
version of a known human GPCR.

XX
SQ Sequence 1329 BP; 330 A; 313 C; 323 G; 363 T; 0 other;

Query Match 30.8%; Score 1322.6; DB 23; Length 1329;
Best Local Similarity 99.7%; Pred. No. 2.5e-266;
Matches 1329; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 238 atgcagcgcctccaaagtctgtgcggacgcgccttggtgcgtggttcttgcctgcgc 297
|||||

Db 1 atgcagcgcctccaaagtctgtgcggacgcgccttggtgcgtggttcttgcctgcgc 60
|||||

QY 298 ctgtcgcgactctggggagagagagagaggttccgcctgcagggccactccgcttttg 357
|||||

Db 61 ctgtcgcgactctggggagagagagagaggttccgcctgcagggccactccgcttttg 120
|||||

QY 358 caaacgcgacagataatgacgcccacccactaagacaccttatggcccaggttccaacgcc 417
|||||
Db 121 caaacgcgacagataatgacgcccacccactaagacaccttatggcccaggttccaacgcc 180
|||||
QY 418 agtctgggcgcgtctgtggcacctcgcggaggtgcctaaaggagacagagcgcgagatct 477
|||||
Db 181 agtctgggcgcgtctgtggcacctcgcggaggtgcctaaaggagacagagcgcgagatct 240
|||||
QY 478 ccgcacacacacatctccctccctccgcgtgcgaaggaccatcgagataagagacatttc 537
|||||
Db 241 ccgcacacacacatctccctccctccgcgtgcgaaggaccatcgagataagagacatttc 300
|||||
QY 538 aaatacatcaacacgcgttgctctgccttctgtctgtctgctgctgggagatcatcggaactcc 597
|||||
Db 301 aaatacatcaacacgcgttgctctgccttctgtctgtctgctgctgggagatcatcggaactcc 360
|||||
QY 598 acactctgagaattatctacaagaacagtgcatgcgaaacgggtcccaatatcttgatc 657
|||||
Db 361 acactctgagaattatctacaagaacagtgcatgcgaaacgggtcccaatatcttgatc 420
|||||
QY 658 gccagcttgctctgggagacactgctgaacatcgctcaatcgatcatccctcatcaatgctctac 717
|||||
Db 421 gccagcttgctctgggagacactgctgaacatcgctcaatcgatcatccctcatcaatgctctac 480
|||||
QY 718 aagctgctggcagagagactggcatttgagctgagatgtgtaagctggtgcctttcata 777
|||||
Db 481 aagctgctggcagagagactggcatttgagctgagatgtgtaagctggtgcctttcata 540
|||||
QY 778 cagaaagcctccgtgggaactactgctgagctctctatgtgctctgagtagttagacagatat 837
|||||
Db 541 cagaaagcctccgtgggaactactgctgagctctctatgtgctctgagtagttagacagatat 600
|||||
QY 838 cgagctgtgtcttctggagtagaataaaaggaattggggttcccaaaatggacagacagta 897
|||||
Db 601 cgagctgtgtcttctggagtagaataaaaggaattggggttcccaaaatggacagacagta 660
|||||
QY 898 gaaattgtttgatttggtggtctctctgtggttctggtctgctgctgctgctgctgctgct 957
|||||
Db 661 gaaattgtttgatttggtggtctctctgtggttctggtctgctgctgctgctgctgctgct 720
|||||
QY 958 gataataacgagtgagactacaaaggaattatctgcgaactgctgctgctgctgctgctgct 1017
|||||
Db 721 gataataacgagtgagactacaaaggaattatctgcgaactgctgctgctgctgctgctgct 780
|||||
QY 1018 cagaagacagctttcatcagcttttacaagacagacaaagattggtggtgctgctgctgct 1077
|||||
Db 781 cagaagacagctttcatcagcttttacaagacagacaaagattggtggtgctgctgctgct 840
|||||
QY 1078 tattctgcttgccattgcccactcactgcatctttttttatcacactaatgacctgtgaaatg 1137
|||||
Db 841 tattctgcttgccattgcccactcactgcatctttttttatcacactaatgacctgtgaaatg 900
|||||
QY 1138 ttgagaagaaaagtgcgcagatgctttaaataatgatccactaaagcagagacgggaa 1197
|||||
Db 901 ttgagaagaaaagtgcgcagatgctttaaataatgatccactaaagcagagacgggaa 960
|||||
QY 1198 gtggccaaaacgcgtcttttgcctggtcttctgtcttgcctctgctgctgctgctgctgct 1257
|||||
Db 961 gtgagaaaaacgcgtcttttgcctggtcttctgtcttgcctctgctgctgctgctgctgct 1020
|||||
QY 1258 ctacgacgagattctgaagctcactctttataatcagatgatcccaatagatgtgaactt 1317
|||||
Db 1021 ctacgacgagattctgaagctcactctttataatcagatgatcccaatagatgtgaactt 1080
|||||
QY 1318 ttgagcttctgttggtatttgactatatattggtatcaacatggtctcactgaattcctg 1377
|||||
Db 1081 ttgagcttctgttggtatttgactatatattggtatcaacatggtctcactgaattcctg 1140
|||||
QY 1378 attaacccaattgctctgtatttggtgagcaaaagattcaaaaactgctttaagtcatg 1437
|||||
Db 1141 attaacccaattgctctgtatttggtgagcaaaagattcaaaaactgctttaagtcatg 1200
|||||

QY 1438 ttatgctggtgcagtcatttgaagaaaaaacagctcttgaggagaaaagcagtcgtgc 1497
 |||||
 Db 1201 ttatgctggtgcagtcatttgaagaaaaaacagctcttgaggagaaaagcagtcgtgc 1260
 |||||
 QY 1498 taaagttcaaaagcattgatcacggtatgatgacaaacttcggtccagtaataatacagc 1557
 |||||
 Db 1261 taaagttcaaaagcattgatcacggtatgatgacaaacttcggtccagtaataatacagc 1320
 |||||
 QY 1558 tcatcttga 1566
 |||||
 Db 1321 tcatcttga 1329
 |||||

RESULT 13
 AAF21287
 ID AAF21287 standard; DNA; 1578 BP.
 XX
 AC AAF21287;
 XX
 DT 14-MAR-2001 (first entry)
 XX
 XX Human low adenosine antisense oligonucleotide related sequence #2854.
 DE
 XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;
 KW human; airway disorder; bronchoconstriction; lung inflammation;
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;
 KW immunosuppressive; antisthmatic; analgesic; hypotensive; cytotatic;
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200062736-A2.
 XX
 PD 26-OCT-2000.
 XX
 PF 24-MAR-2000; 2000WO-US08020.
 XX
 PR 06-APR-1999; 99US-0127958.
 XX
 PA (UYEC-) UNIV EAST CAROLINA.
 PA (NYCE/) NYCE J W.
 XX
 PI NYCE JW;
 XX
 DR WPI; 2000-679539/66.
 XX
 XX Low adenosine (A) content antisense oligonucleotides which do not
 PT trigger adenosine receptors during metabolism, useful e.g. for treating
 PT cancers and respiratory obstructions -
 PT
 XX Disclosure; Page 1276; 1592pp; English.
 PS
 XX
 CC The present invention describes low adenosine (A) content antisense
 CC oligonucleotides and compositions (I) comprising them. In the antisense
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,
 CC immunosuppressive, antisthmatic, hypotensive and cytotatic activities.
 CC The antisense oligonucleotides and (I) can be used to down-regulate the
 CC expression and/or activity of target polypeptides associated with
 CC lung/respiratory disorders and malignancies, such as stimulating and
 CC activating peptide factors and transmitters, transcription factors,
 CC immunoglobulins and antibodies, antibody receptors, cytokines and
 CC chemokines, endogenously produced specific and non-specific enzymes,
 CC binding proteins, adhesion molecules and their receptors, cytokine and
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central
 CC nervous system (CNS) and peripheral nervous and non-nervous system
 CC receptors, CNS and peripheral nervous and non-nervous system peptide
 CC transmitters, defensins, growth factors, vasoactive peptides and

CC receptors, binding proteins and malignancy associated proteins. The
 CC antisense oligonucleotides may be used in this way to treat disorders
 CC including respiratory obstruction (especially pulmonary obstruction
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies)
 CC and/or surfactant hypoproduction which are associated with a disease or
 CC condition selected from pulmonary vasoconstriction, inflammation,
 CC allergies, asthma, impeded respiration, respiratory distress syndrome
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide
 CC fragments and antisense oligonucleotides used in the exemplification of
 CC the present invention.
 XX

Sequence 1578 BP; 380 A; 378 C; 397 G; 423 T; 0 other;

Query Match 28.4%; Score 1220.4; DB 21; Length 1578;
 Best Local Similarity 99.3%; Pred. No. 5.5e-245;
 Matches 1247; Conservative 0; Mismatches 6; Indels 3; Gaps 2;

QY 178 tgaacttgctgctgaaactgcggagcgccaccgcgacccctcttgagcagtagcagc 237
 || |||||
 Db 177 tgcctcagctcgtgaaactgcg--gcggccaccgcagc-ctctcgagcagtagcagc 233
 |||||
 QY 238 atgcagccgctccaaagtctgtgcgacgcgcctcgtgtgctggtcttctgctcgccg 297
 |||||
 Db 234 atgcagccgctccaaagtctgtgcgacgcgcctcgtgtgctggtcttctgctcgccg 293
 |||||
 QY 298 ctgtcgcggtatctgggagagagagagagcttccgcctcgcagagccactcgcgtttg 357
 |||||
 Db 294 ctgtcgcggtatctgggagagagagagagcttccgcctcgcagagccactcgcgtttg 353
 |||||
 QY 358 caaacgcgagagataatgacgcccaccactaaacacttatggcccaagggttccaaagcc 417
 |||||
 Db 354 caaacgcgagagataatgacgcccaccactaaacacttatggcccaagggttccaaagcc 413
 |||||
 QY 418 agtctggcggtgcgttggcaactgcggaggtgcctaaaggagagcagcgagatct 477
 |||||
 Db 414 agtctggcggtgcgttggcaactgcggaggtgcctaaaggagagcagcgagatct 473
 |||||
 QY 478 cgcgcacgcacacatctccctccctccctcccaaggaccatcgagatcaaggagacttc 537
 |||||
 Db 474 cgcgcacgcacacatctccctccctccctcccaaggaccatcgagatcaaggagacttc 533
 |||||
 QY 538 aaatacatcaaacacggttgcctcctccttgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 597
 |||||
 Db 534 aaatacatcaaacacggttgcctcctccttgcgtgcgtgcgtgcgtgcgtgcgtgcgtgc 593
 |||||
 QY 598 acactctgagaaattatctacaagaacagtgcatcgcaaaacgggtcccaatatcttgatc 657
 |||||
 Db 594 acactctgagaaattatctacaagaacagtgcatcgcaaaacgggtcccaatatcttgatc 653
 |||||
 QY 658 gccagcttggtcctggagacctgcgcacatcgctcatgttgacatccctatcaatgctctac 717
 |||||
 Db 654 gccagcttggtcctgggagacctgcgcacatcgctcatgttgacatccctatcaatgctctac 713
 |||||
 QY 718 aagctcgtgcgagacgactggccatttgagctgagatgtgtaagctgggtgcctttcata 777
 |||||
 Db 714 aagctcgtgcgagacgactggccatttgagctgagatgtgtaagctgggtgcctttcata 773
 |||||
 QY 778 cagaaagcctccgtgggaatcactgtcgtgagtcctatgtcgtcgtcgtcgtcgtcgtcgtcgt 837
 |||||
 Db 774 cagaaagcctccgtgggaatcactgtcgtgagtcctatgtcgtcgtcgtcgtcgtcgtcgtcgt 833
 |||||
 QY 838 cgagctgctgctctctcgtgagtagaataaagggaatcgggggtcccaaaaggaacagcagta 897
 |||||
 Db 834 cgagctgctgctctcgtgagtagaataaagggaatcgggggtcccaaaaggaacagcagta 893
 |||||
 QY 898 gaaattgttttatttgggtgcgtcctcgtggttctggtcgtcgtcgtcgtcgtcgtcgtcgtt 957
 |||||
 Db 894 gaaattgttttatttgggtgcgtcctcgtggttctggtcgtcgtcgtcgtcgtcgtcgtt 953
 |||||

QY 958 gatataattacgatgactacaaaggaagtattctcgcgaatctgctgttcttcaatccgctt 1017
 Db 954 gataaattacgatgactacaaaggaagtattctcgcgaatctgctgttcttcaatccgctt 1013
 QY 1018 cagaagacagcttctcagttttacagacagcaaaagattggtggttctcagtttc 1077
 Db 1014 cagaagacagcttctcagttttacagacagcaaaagattggtggttctcagtttc 1073
 QY 1078 tattctgcttgccattggtccatcactgcatctttttttatcacataatgacctgtgaaatg 1137
 Db 1074 tattctgcttgccattggtccatcactgcatctttttttatcacataatgacctgtgaaatg 1133
 QY 1138 ttgagaaagaaagtggcagtcagagattggttttaaatgatcatcaataagcagagcgggaa 1197
 Db 1134 ttgagaaagaaagtggcagtcagagattggttttaaatgatcatcaataagcagagcgggaa 1193
 QY 1198 gtggccaaaacgtcttttgcctggtccttctgtcttcttgcctctgctggttcccttcac 1257
 Db 1194 gtggccaaaacgtcttttgcctggtccttctgtcttcttgcctctgctggttcccttcac 1253
 QY 1258 ctacagcagattctgaagctcactcttttataatcagaatgatcccaatagatgtaactt 1317
 Db 1254 ctacagcagattctgaagctcactcttttataatcagaatgatcccaatagatgtaactt 1313
 QY 1318 ttgagcttcttctgttattggaactatattggtatcaacatggttctcactgaattctgc 1377
 Db 1314 ttgagcttcttctgttattggaactatattggtatcaacatggttctcactgaattctgc 1373
 QY 1378 attaaccaattgctgtattgttgtagcaaaaagattcaaaaactgttttaagtc 1433
 Db 1374 attaaccaattgctgtattgttgtagcaaaaagattcaaaaactgttttaagtc 1429

RESULT 14

AAA35165

ID AAA35165 standard; DNA; 1578 BP.

XX AC AAA35165;

XX DT 28-JUL-2000 (first entry)

XX DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:39.

XX KW Human; adenosine receptor; low adenosine antisense oligonucleotide;
 KW phosphorothioate; impaired respiration; inflammation; allergy;
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;
 KW antiallergic; antiasthmatic; cytostatic; analgesic; impaired airway;
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.

XX OS Homo sapiens.

XX PN WO200009525-A2.

XX PD 24-FEB-2000.

XX PF 03-AUG-1999; 99WO-US17712.

XX PR 03-AUG-1998; 98US-0095212.

XX PA (UYEC-) UNIV EAST CAROLINA.

XX PI Nyce JW;

XX DR WPI; 2000-205971/18.

XX PT New antisense oligonucleotides useful for treating e.g. pulmonary
 PT vasoconstriction, inflammation, allergies, asthma, hypertension, or
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or
 PT cancers -

XX

PS Disclosure; Page 1194; 1343pp; English.

XX CC The present invention describes a new composition comprising an
 CC antisense oligonucleotide (ON) with low adenosine (up to 15%), which
 CC targets nucleic acids involved in bronchoconstriction, allergies, and/or
 CC inflammation. The ON can have antiinflammatory, antiallergic,
 CC antiasthmatic, cytostatic and analgesic activities. The compositions are
 CC useful for the treatment of diseases associated with inflammation,
 CC impaired airways, including lung disease and diseases whose secondary
 CC effects afflict the lungs of a subject. They can be used for treating
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies,
 CC asthma, impaired respiration, respiratory distress syndrome, pain, cystic
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive
 CC pulmonary disease (COPD), and cancers such as leukemias, lymphomas,
 CC carcinomas, and cancers which may metastasize to the lungs, including
 CC breast and prostate cancer. The reduction of the adenosine content of
 CC the ONs reduces side effects. The A-containing ONs break down with the
 CC release of deoxyadenosine which activates adenosine receptors causing
 CC bronchoconstriction and inflammation. AAA32313 to AAA3512 represent the
 CC nucleotide sequences given in the sequence listing from the present
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last
 CC 185 sequences are also called SEQ ID NO:1 to 185, but the sequences
 CC differ from the previously named sequences. SEQ ID NO:11 to 1880
 CC (AAA32323 to AAA33992) are specifically claimed ONs from the present
 CC invention. N.B. Sequences given in the disclosure of the present
 CC invention do not match up with their corresponding SEQ ID NO: sequences
 CC given in the sequence listing.

XX SQ Sequence 1578 BP; 380 A; 378 C; 397 G; 423 T; 0 other;

Query Match

Best Local Similarity 28.4%; Score 1220.4; DB 21; Length 1578;

Matches 1247; Conservative 99.3%; Pred. No. 5.5e-245;

Mismatches 0; Indels 6; Gaps 2;

QY 178 tgaacttggtctgaaactggagcgccacccgacgcttctcttgagcaggtagcgc 237
 Db 177 tgtctctaggtctgaaactgc--gcggccacgcgacg--cttctggagcaggtagcgc 233
 QY 238 atgcagccgcctccaaagtctgtgcgacgcgccttggtctgctgttctgcctgcgc 297
 Db 234 atgcagccgcctccaaagtctgtgcgacgcgccttggtctgctgttctgcctgcgc 293
 QY 298 ctgtcgcgcatctgggagagagagagccttccgcctgacagggccactccgctttg 357
 Db 294 ctgtcgcgcatctgggagagagagagccttccgcctgacagggccactccgctttg 353
 QY 358 caaacgcgagataatgatgcgccaccactaagaccttatggcccaagggttccaaagcc 417
 Db 354 caaacgcgagataatgatgcgccaccactaagaccttatggcccaagggttccaaagcc 413
 QY 418 agtctggcgcgctgctgtgcacactgcggaggtgcctaaaggagacagcgaggatct 477
 Db 414 agtctggcgcgctgctgtgcacactgcggaggtgcctaaaggagacagcgaggatct 473
 QY 478 ccgcacgcacacatctccctcccggtcccaaggaccatcgagatcaaggagacttc 537
 Db 474 ccgcacgcacacatctccctcccggtcccaaggaccatcgagatcaaggagacttc 533
 QY 538 aaatacatcaaacaggtgtgtctctccttctgttctgtctgtctgtctgtctgtctgtct 597
 Db 534 aaatacatcaaacaggtgtgtctctccttctgttctgtctgtctgtctgtctgtctgtct 593
 QY 598 acacttctgagaattatctacaagaacaagtgcacgaaacaggttcccaatatctgac 657
 Db 594 acacttctgagaattatctacaagaacaagtgcacgaaacaggttcccaatatctgac 653
 QY 658 gccagcttgctctgggagacactgtgcacatcgctcatcgatgacatccctatcaatgctac 717
 Db 654 gccagcttgctctgggagacactgtgcacatcgctcatcgatgacatccctatcaatgctac 713
 QY 718 aagctgctggcagagagactggccatttggagctgagatgtgtgaagctgggttccttcata 777
 Db 718 aagctgctggcagagagactggccatttggagctgagatgtgtgaagctgggttccttcata 777

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 1, 2002, 00:41:00 : Search time 14667 Seconds
(without alignments)
6136.584 Million cell updates/sec

Title: US-09-931-157-2

Perfect score: 4301

Sequence: 1 gagacattccggtggggac.....ctgggaaaaaaaaaaaaa 4301

Scoring table:

IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues

Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

1: gb_ba:*

2: gb_hgt:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_hgt_hum:*

31: em_hgt_inv:*

32: em_hgt_other:*

33: em_hgtgo_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description

Query					

1	4301	100.0	4301	6	AR177880	AR177880 Sequence
2	4301	100.0	4301	6	E07650	E07650 cDNA encodi
3	4286	99.7	4286	9	HUMETR	D90402 Human mRNA
4	4284	99.6	4286	9	S57283	S57283 endothelin
5	2857	66.4	2972	9	HUMHETBR7	D13168 Human gene
6	2841.8	66.1	183337	9	AL139002	AL139002 Human DNA
7	2610	60.7	2720	11	G06417	G06417 human STS W
8	1691.8	39.3	1873	6	AR165435	AR165435 Sequence
9	1691.8	39.3	1873	6	E15242	E15242 Human mRNA
10	1690.8	39.3	1872	9	S44866	S44866 ETB endothe
11	1495.4	34.8	1719	9	HUMEDNRB	L06623 Homo sapien
12	1466.8	34.1	1470	9	HUMETSR	M74921 Human endot
13	1389	32.3	1603	9	BC014472	BC014472 Homo sapi
14	1399	32.3	1632	6	AX342673	AX342673 Sequence
15	1361.4	31.7	1765	9	AF114165	AF114165 Homo sapi
16	1322.6	30.8	1329	6	AX280873	AX280873 Sequence
17	1222.8	28.4	1669	4	AF019072	AF019072 Equus cab
18	1220.4	28.4	1578	9	HSX99250	X99250 H.sapiens m
19	1197.8	27.8	2026	4	BOVEETR	D90456 Bovine mRNA
20	1186	27.6	1452	4	AF034530	AF034530 Canis fam
21	1113	25.9	2018	10	S65355	S65355 nonselectiv
22	1110.2	25.8	1851	6	E05930	E05930 DNA sequence
23	1099.2	25.6	1958	6	AX305434	AX305434 Sequence
24	1099.2	25.6	1958	10	MMU32329	U32329 Mus musculu
25	1091	25.4	1892	10	RNETBREC	X57764 Rat mRNA fo
26	1091	25.4	1965	6	E03623	E03623 DNA encodin
27	1086.6	25.3	1311	4	AF038900	AF038900 Equus cab
28	1067.6	24.8	1314	4	AF276427	AF276427 Canis fam
29	1042.8	24.2	1326	4	AF245459	AF245459 Oryctolag
30	732.6	17.0	1041	5	CCEDNRB	X99295 C.coturnix
31	588	13.7	588	11	G15922	G15922 human STS C
32	564.8	13.1	1520	5	AF275636	AF275636 Danio rer
33	538	12.5	1724	5	AB045356	AB045356 Oryzias l
34	537.2	12.5	1752	9	HUMHETBR1	D13162 Human gene
35	504	11.7	1308	5	CCEDNRB2	Y16089 Coturnix co
36	440.6	10.2	592	6	AX331475	AX331475 Sequence
37	440.6	10.2	592	6	AX332481	AX332481 Sequence
38	427.6	9.9	1436	10	RATENDOR	M60786 Rat endothe
39	414.8	9.6	1180	4	AF311974	AF311974 Oryctolag
40	414.4	9.6	2944	5	AF040634	AF040634 Gallus ga
41	413.2	9.6	2696	10	BC008277	BC008277 Mus muscu
42	411.8	9.6	1374	4	S80652	S80652 endothelin
43	411.8	9.6	3216	4	BTBETREC	X57765 Bovine mRNA
44	401.8	9.3	1650	5	XLU06633	U06633 Xenopus lae
45	399	9.3	1284	4	AF416703	AF416703 Ovis arie

ALIGNMENTS

RESULT	1					
AR177880						
LOCUS	AR177880	4301 bp	DNA	linear	PAT 17-DEC-2001	
DEFINITION	Sequence 3 from patent US 6313276.					
ACCESSION	AR177880					
VERSION	AR177880.1					
KEYWORDS	GI:17920235					
SOURCE	Unknown.					
ORGANISM	Unknown.					
REFERENCE	Unclassified.					
AUTHORS	1 (bases 1 to 4301)					
TITLE	Imura, H., Nakao, K. and Nakanishi, S.					
JOURNAL	Human endothelin receptor					
FEATURES	Patent: US 6313276-A 3 06-NOV-2001;					
	Location/Qualifiers					
	source					
	1. .4301					
	/organism="unknown"					
BASE COUNT	1342 a	828 c	817 g	1314 t		
ORIGIN						

Query Match 100.0%; Score 4301; DB 6; Length 4301;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 4301; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy 1 gagacattccggtggggaactctggccagcccgagcaactggatctctgagcaactccc 60
Db 1 GAGACATTCGGGTGGGGGACTCTGGCCAGCCGAGCAACGTGGATCTCGAGACACTCCC 60

Qy 61 agtaggcatcttcccggtgggagcgccttcgcaagaacagtgtgtgacagcccccggtg 120
Db 61 AGTAGGCAATTCGCCCGGTGGGACGCTTGGCCAGACAGCTGTGTGGCAGGCCCGCGTGG 120

Qy 121 aggatcaacacagtggctgaacactgggaggaactggtacttgagcttgagcttgacatcga 180
Db 121 AGGATCAACACAGTGGCTGAACACTGGGAAGGAACCTGGTACTTGGAGTCTGGACATCTGA 180

Qy 181 aacttggtctgaactgcggagcggccaccgcgacgcctcttcggagcaggtagcagcatg 240
Db 181 AACTTGGCTCTGAACATGCGGAGCGGCCACCGGACGCTTCTGGAGCAGGTAGCAGCATG 240

Qy 241 cagccgctccaagtctgtgcgagcgcgcctggctggctggtcttgcctgctggcctg 300
Db 241 CAGCCGCCCTCCAAGTCTGTGCGGACGCGCCCTGGTTGCGTGGTTCCTGTGCGGCGCTG 300

Qy 301 tcgcgatctcgggagagagagagcttcccgctgaacagggccactccgcttttgcaa 360
Db 301 TCCGGATCTGGGAGAGAGAGAGGCTTCCCGCTGACAGGGCCACTCCGCTTTTGCAA 360

Qy 361 accgcagagataatgacgcgccaccactaagaccttatggccaaagggttccaacgcaggt 420
Db 361 ACCGCAGAGATATGACGCCACCCTAAGACCTTATGGCCCAAGGGTTCCAACGCCAGT 420

Qy 421 ctggcggtcgtgttgccacctcggagggtccttaaggagacaggaagcaggtatcccg 480
Db 421 CTGGCGGTCGTGTGGCACCCTGGGAGGTGCTTAAGGAGACAGGACGCGGAGGATCTCCG 480

Qy 481 ccaagcaccatctccctccctccctgcccagagaccatcgagatcaaggagactttcaaa 540
Db 481 CCAGCACCACATCTCCCTCCCGTCCCAAGGACCCATCGAGATCAAGGAGACTTTTCAA 540

Qy 541 tacatcaacacaggtgtgtgctcgtccttctgtctgctggggatcatcgggaactccaca 600
Db 541 TACATCAACACGGTTGTGCTGCTGCTTGTGCTGCTGGGATCATCGGNACTCCACA 600

Qy 601 cttctgagaattctacagaacaaagtgcgtaagcgggtcccaaatctcttgatgcc 660
Db 601 CTTCTGAGAATTATCTACAGAACAAGTGCATCGAAGCGGTCCCAATATCTTGATCGCC 660

Qy 661 agcttggtctgagagacactgcgcacatcgtcattgacatccctatcaatgctacaag 720
Db 661 AGCTTGGCTCTGGGAGACCTGCTGCACATCGTCAATTGACATGCTTATCAATGTCTACAAG 720

Qy 721 ctgctggcagagactggccatttgagctgagatgtgtgaagctggctttcatcacag 780
Db 721 CTGCTGGCAGAGACTTGGCCATTTGGAGCTGAGATGTGTAAAGCTGGTCCCTTTTCATACAG 780

Qy 781 aaagccctcgtgggaatcactgctgagctctatgtgctgctgagtatggagatatcga 840
Db 781 AAAGCCCTCGTGGAATCACTGTGCTGAGTCTATGTGCTGCTGAGTATTGACAGATATCGA 840

Qy 841 gctgtgtcttcttgagtagaataaagaattcggggtcccaaaatggacagcagtagaa 900
Db 841 GCTGTGTCTTCTTGGAGTAGAATTAAGGAATTTGGGTTCCAAAATGGACAGCAGTAGAA 900

Qy 901 attgtttgattgggtgctcgtggttctggtctgctgacatccctgaaccataggtttgat 960
Db 901 ATTGTTTTGATTTGGGTGGTCTCTGTGGTCTGCGCTGCTCCCTGAAGCCATAGGTTTGTAT 960

Qy 961 ataattacgatggactacaagaagatctatctcgcaatcgttctgcttcacccgttcag 1020
Db 961 ATAATTACGATGGACTACAAGAAAGTATCTCGGAATCTGCTTGTCTATCCCGTTCAG 1020

Qy 1021 aagacagcttctaatcagttttacaagacagcaaaagattggtggtgttcagtttctat 1080
Db 1021 AAGACAGCTTTTCATGCAGTTTTACAAGACAGCAAAAGATTGTTGGTGTTCAGTTTCTAT 1080
```

```
Qy 1081 ttctgcttgccattggccatcactgcattttttatatacactaatgacctgtgaaattgtg 1140
Db 1081 TTCtGCTTGCCATTGGCCATCAGTGCATTTTATATACATTAATGACCTGTGAATGTTG 1140

Qy 1141 agaaagaaaagtggcatgcagattgctttaaatgatacactaaagacagagacggaagtg 1200
Db 1141 AGAAAGAAAAGTGGCATGCAGATTGCTTTAAATGATACACTAAAGCAGAGACGGGAAGTG 1200

Qy 1201 gccaaaaccgctcttggctggtccttcttggcctctcctgctgcttccccttccacctc 1260
Db 1201 GCCAAAACCGCTTTTGGCTGGTCTTGTCTTGTGCCCCTGCTGCTGCTCCCTTCACCTTC 1260

Qy 1261 agcaggattctgaagctcactctttataatcagaatgatcccaatagatgtagaacttttg 1320
Db 1261 AGCAGGATTCGAAGCTCAGCTCTTTATAATCAGAATGATCCCAATAGATGTAACATTTTG 1320

Qy 1321 agcttctgttggtattggaactatattggtatcaacatcggtctcactgaattccgtcatt 1380
Db 1321 AGCTTCTGTGTTGGTATTGGACTATATTGCTATCAACATGGGTTCTACTGAATTCCTGCATT 1380

Qy 1381 aacccaattgctctgtatttggtagcaaaaagattcaaaaactgctttaaagtcagctta 1440
Db 1381 AACCCAATTGCTCTGTATTGTTGGTGAAGAAAAGATTCAAAAAGTCTTAAAGTCATGCTTA 1440

Qy 1441 tgctgctgggtccagtcatttggaagaaaaacagtcctcttgaggaaaaagcagctgctta 1500
Db 1441 TGCTGCTGGTGGCCAGTCAATTTGAAGAAAACAGTCTTGGAGGAAAACAGCTGCTGCTTA 1500

Qy 1501 aagttcaagctaatgatcacggatgacagacttccgttccagtaataataacagctca 1560
Db 1501 AAGTTCAAAGCTAATGATCAGCGATATGACAACTTCCGTTCCAGTAATAAATACAGCTCA 1560

Qy 1561 tcttgaagaagaactattcactgtatttcttcttcttcttcttcttcttcttcttcttctt 1620
Db 1561 TCTTGAAGAAGAAGTATTTCATCTGTATTTCATTTCTTTATATTGGACCGAAGTCATTAA 1620

Qy 1621 acaaaaatgaacatttggccaaaaacaaacaaaactatgtatttgcacagcacactat 1680
Db 1621 AACAAAATGAACATTTGGCCAAAACAAAACAAAACATATGATTTGCACAGCACACTAT 1680

Qy 1681 taaaattaaagtgaattattttaacactcacagctacatgatgacattttatgagctgt 1740
Db 1681 TAAATATTAAAGTAAATTTATTTAACTCACAGTACATGATGATGATGATGATGATGATG 1740

Qy 1741 ttacggcatggaaagaaaaatcagtggaattaaagaagcctcgtcgtgtaaaagcacttaat 1800
Db 1741 TTACGGCATGGAAAGAAAATCAGTGGGAATTAAGAAAGCCCTCGTGTGAAAGCAGCTTAAT 1800

Qy 1801 tttttacagtttagcacttcaacatagctcttcaacaaactccagagatatccacacacact 1860
Db 1801 TTTTACAGTTAGCAGCTTCAACATAGCTCTTTAACAACTTCCAGGATATTTCACACAACT 1860

Qy 1861 taggcttaaaatgagctcacgcagaatttctattcttcttaaaagagatttttttta 1920
Db 1861 TAGGCTTTAAAAATCAGCTCAGTCAAGAAATTTCTATTCTTTCTTTAAAAAGAGATTTATTTTA 1920

Qy 1921 aatcaatgggactctgataaaaggaagaaataagtcactgttaaaacagaaacttttaaatg 1980
Db 1921 AATCAATGGGACTCTGATATAAGGAAGAAATAAGTACTGTATAAACAGAACTTTTAAATG 1980

Qy 1981 aagcttaaaatctcaatttaaaattttaaactctttaaacaacttttcaataaat 2040
Db 1981 AAGCTTAAATCTCAATTTTAAATTTTAAATTTTAAATTTTAAACAACTTTTCAATTAATAT 2040

Qy 2041 taccacactattacagattgtaattagatgcgaatgagagagcaggttagttgttgcatt 2100
Db 2041 TATCACACTATTATCAGATTGTATTAGATGCAATAGAGAGAGCAGTTTAGTTGTGCAT 2100

Qy 2101 ttttcgacactggaaacatttaaatgatcagggaggagtagtaacagaaagagcaggtcgt 2160
Db 2101 TTTTCGACACTGGAACATTTAAATGATCAGGAGGGAGTAAACAGAAAGCAAGGCTGT 2160
```

Qy	2161	ttttgaaatacattacactttcactagagccccaaacctcagcatctcgtcaatatgtaac	2220
Db	2161	TTTTGAAATCATTTACACTTTCTACTAGAGGCCAAACCTCAGCATCTCGCAATATGTAA	2220
Qy	2221	caacatgtcaacaacaagcagcatgtaacacagctggcagacatgtgccagctgaatttaaaa	2280
Db	2221	CAACATGTCAACAACAACGACGATGTACAGACTGGCCACATGTGGCAGCTGAATTTAA	2280
Qy	2281	tataatacttttaaaagaaaaattattacatccttttacattcagtttaagatcaacaacctca	2340
Db	2281	TATAATACTTTTAAAGAAAAATTTATACATCTTTTACATTCAGTTAGATCAAACTTCA	2340
Qy	2341	caagagaaaatagaaatgtttgaaaggctatccccaaagaccttttttgaaatgttcattca	2400
Db	2341	CAAGAGAAAATAGAAATGTTTGAAGGCTATCCCAAAAGACTTTTGTGAATCTGTCAATTC	2400
Qy	2401	catacctgtgaagacaatactactctcaaattttttcaaggattattaaaaactctcttttt	2460
Db	2401	CATACCTGTGAAGACAATACTACTCAAAATTTTTCAGGATTTATAAAACTTCTCTTTT	2460
Qy	2461	tcactatcgtatgacttaaaactcgtttggtttctcactctgtaaatcaacttacctacataca	2520
Db	2461	TCACTATCTAGCTTAACTCTGTTGGTTTGTCTCATCTGTAATACTTACCTACATACA	2520
Qy	2521	ctgcatgtagatgattaaatgaggcgagccctgtgctcatagctttacgatggagagat	2580
Db	2521	CTGCATGTAGATGATTAATCAGGCGAGCCCTGTGCTCATAGCTTTACGATGGAGAGAT	2580
Qy	2581	gcaagtgaacctcaataaagaactgtgaactgacctgctggtgcagtgctccacatgcaaaaggg	2640
Db	2581	GCCAGTGACCTCAATAAAGACGTGAACGTGCTGTGTCAGTGTCCACATGCACAAGGG	2640
Qy	2641	gcaggtgagccctctctcacccctcgtcgtggtttaaagtgttctcagcatatataat	2700
Db	2641	GCAGGTAGCCACCTCTCTCACCCATGTGTGGTTTAAATGGTTTCTAGCATATGTATAT	2700
Qy	2701	gctatagttaaaatactatttttcaaaatacacagattagcacatttaacagctacctg	2760
Db	2701	GCTATAGTTAAAATACTATTTTCAAAATCATACAGATTTAGTACATTTAACAGCTACCTG	2760
Qy	2761	taaagcttattactaaatttttgtattattttgttaaatagccaatagaaaaagtgtgcttg	2820
Db	2761	TAAAGCTTATTACTAAATTTTGTATATTTTGTAAATAGCCAATAGAAAAGTTGTCTG	2820
Qy	2821	acatggtgctttcttccatcagagcgaactcgtttttagaccgtgaagaacctctt	2880
Db	2821	ACATGGTGCTTTCTTCTCATCTAGAGCCAAACTGCTTTTTCAGACCCGTAGAACCTCTT	2880
Qy	2881	agcttbtgogtctcctgctaatttttatcttcttaagcaaatgcttaggatgctt	2940
Db	2881	AGCTTTGTGCTTCTGCTCAATTTTATATCTTCTTAAGCAAAAGTGCTTAGGATAGCTT	2940
Qy	2941	ggagtagatgtgtgaaagtatgtacaagagaaaaacggaagagaggaatgaagtg	3000
Db	2941	GGATGAGATGTGTGTAAGATATGTACAGAGAAAACGGAAGAGAGAGGAAATCAGGTG	3000
Qy	3001	gggttgaggaaaacccatgggacagatctccattcttagccctcagcttcgtcattgcct	3060
Db	3001	GGGTTGGAGGNAACCCATGGGACAGATTCCTCATCTTAGCCTAACGTTCTGTCATGCCT	3060
Qy	3061	cgtcacatcaatgcaaaagtcctgattttgttccagcaaaaacacagctgcaatgtctca	3120
Db	3061	CGTCACATCAATGCAAAAGTCTGTATTTGTTCCAGCAAAACACAGTGCNAATGTTCTCA	3120
Qy	3121	gagtgacttcgaaataaattggggccaaagactttaactcgtgctttaaataatgcctcaa	3180
Db	3121	GAGTGACTTTCGAATAAATGGGCCCAAGAGCTTTTAACTCGGTCTTAAAAATATGCCAA	3180
Qy	3181	atttttactttgtttcttttaatagcctggggccacatgtttgaaataagctagtaatg	3240
Db	3181	ATTTTTACTTTGTGTTTTCTTTTAAAGGCTGGGCCACATGTTGGAAAATAGCTAGTAATG	3240
Qy	3241	ttgtttctgtccaatatgtgaatgtgaatgtaacataaaccaaaccaacaataatgtgccca	3300
Db	3241	TTGTTTTCTGTCAATATTGAATGTGATGTTACAGTAACCAAAACCAACAAATGTGGCCA	3300
Qy	3301	gaaagaaagacaataaataaattcaacacacacatatggattcttattataaatcaacc	3360
Db	3301	GAAAGAAAGACAATAATAATTCAATTCACACACCATATGGATTTCTATTATAAATCACC	3360
Qy	3361	acaaacttgttctttaatttcatcccaatcacttttccagagcctgttatcatagaagt	3420
Db	3361	ACAAACTTGTCTTTTAATTTTCATCCCAATCACTTTTTCAGAGGCTGTTATCATAGAAGT	3420
Qy	3421	cattttagactctcaaattttaaaatttttgaatcacataatattttcacagtttatttaa	3480
Db	3421	CATTTTAGACTCTCAATTTTAAATTTTGAATCACTAATATTTTACAGTTTATTATAA	3480
Qy	3481	tatatattaattctatttaaaatttttagatattttttatcacatgtaactgtttttaca	3540
Db	3481	TATATTTAATTTCTATTTTAAATTTTAGATTATTTTATTACATGTACTGAATTTTACA	3540
Qy	3541	tccgtgatacccttctctcctcagtcagtcagtcagtcctcaattcttctgcgaatttt	3600
Db	3541	TCCTGATACCCCTTTCCTTCTCATGTCAAGTATCAATGTTCTTAATTTATCTTGCCAAATTT	3600
Qy	3601	tgaactcacacacaaaaagcactctgcatattttataataaaattgcatcagtggttt	3660
Db	3601	TGAAACTACACACAAAAAGCATCTTGTCATTTATTATAATAAATTCATTCAGTGCTT	3660
Qy	3661	tttaaaaaaaatgtttgattcaaaactttcaactactgataagtaagaaaaaatataat	3720
Db	3661	TTTAAAAAAATGTTGATTCAAACTTTAACTACTGATGAAGTAAGAAACAATTTATAT	3720
Qy	3721	tctttcatactcaaaaccaaagtagaaaaagtgctcagtcagtcacacttcaaacatgt	3780
Db	3721	TTCTTTACATACTCAAAACCAAGATAGAAAAGGTGCTATCTTCAACTTCAAAACATGT	3780
Qy	3781	ttcctagtattaaaggactttaatatagcaacagacaaaattattgttaacctggatgta	3840
Db	3781	TTCTAGTATTAAAGACTTTTAAATAGCAACAGACAAAAATTTGTTTAACTGGATGTTA	3840
Qy	3841	cagctcaaaaagattataaaaagatttaaacctatttttcccttattatccactgctaat	3900
Db	3841	CAGCTCAAAAAGATTTATAAAAAGATTTTAACTATTTTCTCCCTTATTATCCACTGCTAAT	3900
Qy	3901	gtgagtgtatgttcaaacacaccttttagtattgtagcttacaatggtgcaagaataca	3960
Db	3901	GTGGATGTATGTTCAACACCTTTTGTATTGTATGATAGTTACATATGCCCAAGGAATACA	3960
Qy	3961	gtttatagcaaaaacatggggtatgctgtagctaaactttataaagtgtataatacaatgt	4020
Db	3961	GTTTATAGCAAAACATGGGTATGCTGTAGCTAACTTTTATAAAGTGAATATAACAATGT	4020
Qy	4021	aaaaaattatatctctgggagattttttgggtgcctaaagcgctactagttactgattt	4080
Db	4021	AAAAAATTTATATCTCTGGAGATTTTGTGGTTCCTTAAAGTGGCTACTAGTTACTGATTT	4080
Qy	4081	tttattatgaagcaaaaccaaataaaatttaaagtttttttcaaacactcacttattttc	4140
Db	4081	TTTATATTATGAAGCAAAACCAATAAAAATTTTAACTTTTATTAACACTTACTCTTTTTC	4140

Qy	1381	aaccaatctgctctgtatttggtagcaaaaaagattcaaaaactcgttttaagtcattgctta	1444
Db	1381	AACCAATTCGCTGCTATTTGGTGTGACAAAAAGATTCAAAAACTCGCTTTTAAGTCAATGCTTA	1440
Qy	1441	tgcctgcgtgcgcagctcatttgaaaaaaacagctccttggaggaaaaagcagtcgtgctta	1500
Db	1441	TGCTGCTGGTGGCCAGCTCATTTGAAGAAAACAGTCTTGGAGGAAAACAGCTCGTGTCTTA	1500
Qy	1501	aagttcaaaagtaaatgatcacgagatagacaacttccgttccagtaaataatcacagctca	1560
Db	1501	AAGTTCAAAGCTAATGATCACGSGATATGACAACCTCCGTTCCAGTAATAAATACAGCTCA	1560
Qy	1561	tcttgaagaagaactattcactgtatttcatttcttcttatattggaccggaagtcattaa	1620
Db	1561	TCCTTGAAGAAGAACTATTCACTGTATTTCATTTCTTTTATATTGGACCGAAGTCATTAA	1620
Qy	1621	aacaaaatgaacatttgcacaaacacaaacaaaacaaactatgatttgcacagcacactat	1680
Db	1621	AACAAAATGAACATTTGCCAAACAAAACAAAACATATGTATTGTGCACAGCACACTAT	1680
Qy	1681	taaaataataagtgtaattattttaacactcacagctcacatattgacatttttatgagctgt	1740
Db	1681	TAAAATATTAAAGTGTAATTATTTAACACTCACAGCTACATATGACATTTTATGAGCTGT	1740
Qy	1741	tacgcgcatggaagaataatcagctgggaaattaagaaagccctcgctgtagaagcactcaat	1800
Db	1741	TTACGGCATGGAAAGAAAATCAGTGGGAATTAAGAAAACCTCGTGTGAAAAGCACTTAAT	1800
Qy	1801	tctttcagcttagcacttcaacatagctcttcaaacactccagatatctcacacaacact	1860
Db	1801	TTTTTACAGTTAGACTTCAACATAGCTCTTAACAACTTCACAGATATTCACACAACACT	1860
Qy	1861	taggcttaaaaatgagctcactcagaattcttattctttcttaaaaagagatttatatttta	1920
Db	1861	TAGGCTTAAAAATAGCTCAGTCTCAGAAATTTCTATTCTTTTCAAAAAGAGATTATTTT	1920
Qy	1921	aatcaatgggaactctgataaagaagaataagtcactgtgtaaacagagaacttttaaatg	1980
Db	1921	AATCAATGGGACTCTGATATAAGGAAGAATAAGTCACTGTAAAACAGAACTTTTAATG	1980
Qy	1981	aagcttaaatctcaatttcaaaatttcaaaatcctttcaaacacttttcaatttaatat	2040
Db	1981	AAGCTTAAATTACTCAATTTTAAATTTTAAAATCTCTTTAAAACAACCTTTTCAATTAATAT	2040
Qy	2041	tatacactattatcagattgtaattagatgcataatgagagagagcagtttagttgctgcat	2100
Db	2041	TATCACACTATTATCAGATTGTAATTAGATGCAANTGAGAGAGCAGTTTACTTGTTCGAT	2100
Qy	2101	ttttggacactggaacaatttaaatgatcaggaggaggtaacacagaagagcaggctgt	2160
Db	2101	TTTTTCGGACACTGGAAACATTTAAATGATPCAGAGGGGAGTAACAGAAGAGACAAGGCTGT	2160
Qy	2161	ttttgaaatcattacactttcactagaagcccaaacctcagcattctgcaatatgtaac	2220
Db	2161	TTTTTGAAATTCATTACACTTTTCACTAGAAGCCCAACCTCAGCATTTGCAANTATGTAAAC	2220
Qy	2221	caacatgtcacaaacaagcagcatgttaacagactggccacatgtgccagctgaaatttaaaa	2280
Db	2221	CAACATGTCACAACAGCAGCATGTACAGACTGGCACATGTGCCAGCTGAATTTAAAA	2280
Qy	2281	tataatacttttaaaagaataattattacatcctttacattcagtttaagatcaaacctca	2340
Db	2281	TATAAATCTTTTAAAAAGAAATTTATTACATCTCTTTACATTTTACATTTAAGATCAAAACCTCA	2340
Qy	2341	caaaaggaataagaatgtttgaaaggctatccccaaaagactttttgaaatctgtcatca	2400
Db	2341	CAAAGAGAATAAGAATGTTTGAAGGCTATCCCAAAAGACTTTTTCGAATCTGTCAATCA	2400
Qy	2401	cataccctgtgaagacaatactatctacaaatttttttcaggattattaaaatctctctttt	2460
Db	2401	CATACCCTGTGAAGACAAATACTATCTACAAATTTTTCAGGATTTATTTAAAATCTCTCTTTT	2460
Qy	2461	tcactatcgtagcttaaacctcgtttggtttttgttcactgttaaaatacttaacctacata	2520

D	b	2461	TCACTATCGTAGCTTAAACTCTGTTTGGGTTTTGTCACTGTAAAACACTTACCCTACATACA	2520
Q	y	2521	ctgcatgcagatgatataaatgaggcgagcccttgctcctacatgcttaccatgcagagagat	2580
D	b	2521	CTGCATGTAGATGAATAATGAGGCGAGCCCTTGCTCATAGCTTTACGATGGAGAGAT	2580
Q	y	2581	gccgtgacctcataataaagactgtgaactcctgggtgcagctgccacatgacaaaggg	2640
D	b	2581	GCCAGTCACCTCATAATAAAGACTGTGAAC TGCTGGTGCAGTGTCCACATGACAAAGG	2640
Q	y	2641	gcaggtagcacctctctcccacatgctgtggttaaaatggtttcctagcatagtataat	2700
D	b	2641	GCAGGTAGCACCTCTCTCACCATGCTGTGGTTAAAAATGGTTTTCTAGCATATGTATAT	2700
Q	y	2701	gcatagttaaaaatactattttccaataatcacagattagtacatttaaagcgtaccctg	2760
D	b	2701	GCTATAGTTAAAATACTATTTTTCAANAATCATACAGATTAGTATCAATTTAACAGCTACCTG	2760
Q	y	2761	taaagctattactaatatttgtattattttgtaaaaagccaatagaaaagtttgcttg	2820
D	b	2761	TAAAGCTTATTACTAATTTTGTATTATTTTGTAAATAGCCAATAGAAAAGTTTGCTTG	2820
Q	y	2821	acatggtgctttcttttcactcagaggcaaaactgctttttgagaccgttaagaacctttt	2880
D	b	2821	ACATGGTGCCTTTCTTTCATCTAGAGGCCAAAAC TGCCTTTTGAGACCGTAAAGAACCCTCT	2880
Q	y	2881	agctttgtcgcttcctgcctcaatttttatcttctaagcaaatgaccttaggatagctt	2940
D	b	2881	AGCTTTGTGCGTTCCTGCCCTAAATTTTATCTTTATCTTTAGCAAAGTGCCCTTAGGATAGCTT	2940
Q	y	2941	gggatgagatgctgtgaaagtatgtacaagagaaaaacggaagagagagaaaatgagtg	3000
D	b	2941	GGGATGAGATGCTGTGAAAGTATGTACAAAGAGAAAACGGAAGACAGAGAAATCAGGTG	3000
Q	y	3001	gggttgaggaaaccocatggggagagattccccattcttagcctaactctgtcattgctt	3060
D	b	3001	GGGTTGGAGGAACCCATGGGGACAGATTCCCATTCTTAGCCTAACGTTCTGCTCATGTGCT	3060
Q	y	3061	cgtcacataaatgcaaaaggtcctgatttgttccagcaaaaacacacagtgcaatgtctca	3120
D	b	3061	CGTCACATCAANTGC AAAAGGTCCTGATTTTGTTCAGCANAAAACACAGTGCAATGTTCTCA	3120
Q	y	3121	gagtgacttcgaaaaataattgggcccagaagctttaactcggtcttaaaaatagcccaa	3180
D	b	3121	GAGTGACTTTCGAAAATAAATTTGGGCCCAAGAGCTTTAACTCGGTCTTAAAATATGCCAA	3180
Q	y	3181	attttacttgttttcttttaataaggctggggcacacatgttggaataaagctagttaatg	3240
D	b	3181	ATTTTACTTTGTTTTTCTTTTAATAGGCTGGGCCACATGTTGGAATAATAGCTAGTAATG	3240
Q	y	3241	tgttttctgtcaatttgaatgtgatgtcagtaaacacaaaacccaacaaatgtggcca	3300
D	b	3241	TTGTTTTCTGTCAATATTTGAATGTGATGTACAGTAACCAAMACCCCAACCAATGTGGCCA	3300
Q	y	3301	gaagaaagacaaataataattaattcacacaccataaggattctattataaaatcaccc	3360
D	b	3301	GAAGAAAGACCAATAATAATTTATTCACACACCATATGGATTCTTATTATAAATCACCC	3360
Q	y	3361	acaaactgttctttaaatttcacccaatcaacttttcagaggcctgttatcatagaagt	3420
D	b	3361	ACAAACTGTGTTTAAATTTCAATCCCAATCACCTTTTTCAGAGGCCCTGTATCATAGAAGT	3420
Q	y	3421	caitttagactctcaattttaataatttgaatacactaataattttcaagtttatttaa	3480
D	b	3421	CATTTTAGACTCTCAATTTTAAATTAATTTTGAATCACATAATTTTTCACAGTTTATTTAA	3480
Q	y	3481	tatatatttctattaaatttagattatttttatccatgactgaattttaca	3540
D	b	3481	TATATTTAATTTCTATTTAAAAATTTTATAGATTATTTTATTTACATGTACTGAATTTTTTACA	3540
Q	y	3541	tccgtgaccccttccctctccatgctcagatcatgcttctctctaattatcttgcacaaatt	3600

Db 3541 TCCTGATACCCCTTCCTCTCCATGTCAGTATCATGTTCTCTAATTATCTTGCCAAATTT 3600

Qy 3601 tgaactacacacacacacacactacttgcatatttataataaaattgcattcagtgctt 3660

Db 3601 TGAACACTACACACAAAGACATCTGCATTTATATAATAAAATGCATTCAGTGGCTT 3660

Qy 3661 tttaaaaaaattgtttgattcaaaacttttaacatactgataagtaagaacaaattataat 3720

Db 3661 TTTAAAAAAATGTTTGAATTCAAAACCTTTAAACATCTGATAAGTAAAGAACAAATTAATAAT 3720

Qy 3721 tcttttatactactcaaaacacagatagaaaaagggtgctatcgttcaacttcaaaacatgt 3780

Db 3721 TCTTTTACATCTCAAAACACAGATAGAAAAGGTGCTATCCTTCACTTCAAAACATGT 3780

Qy 3781 tctcagtagtattaaggaactttaataatagcaacagacacaaaattattgttaacatggatgta 3840

Db 3781 TTCTCTAGTATTAAAGGACTTTAATAATAGCAACAGACAAAATTTATTGTAAATGGATGTTA 3840

Qy 3841 cagotcaaaagatttataaaagattttaacctattttctcccttattatccactgctaat 3900

Db 3841 CAGCTCAAAAGATTATATAAGATTTTAAACCTATTATTTCTCCCTTATTATCCACTGCTAAT 3900

Qy 3901 gtggatgtatgttcaaacaccccttttagtattgatgacttacatataggccaaaggaataca 3960

Db 3901 GTGGATGTATGTTCAACACACCTTTTAGTATTGATAGCTTACATATGGCCAAAGGAATACA 3960

Qy 3961 gttatagcaaaaacatgggtatgctgtagctaaactttataaagtgttaatacaaatgt 4020

Db 3961 GTTTATACAAAACATGGTATGCTGTAGCTAACCTTTATAAAGTGTATATAACAATGT 4020

Qy 4021 aaaaaattatatctggtggagatttttgggtgcttaagtggcttagttactgattt 4080

Db 4021 AAAAAATTATATATCTGGGAGGATTTTGGTTGGCTTAAGTGGCTATAGTTACTGATTT 4080

Qy 4081 tttattatgaagcaaaacacataaaatttaagttttttaaacactacottatttttc 4140

Db 4081 TTTATTATGTAAGCAAAACCAATAAAATTTAAGTTTTTTTAAACAACTACCTTATTATTTTC 4140

Qy 4141 actgtacagacactaatcattcaataactaattgattgttttaaaagaataataaagtga 4200

Db 4141 ACTGTACAGACACTAATTCATTAACTAATTAATGATTGTTTAAAGAATAATATAATATGTA 4200

Qy 4201 caatggacattattatgtttaatatatacaattatcaagcaagtagtaagtattcaatt 4260

Db 4201 CAATGGACATTATTATGTTAAATATACAATTATCAAGCAAGTATGAAGTTATTCAATT 4260

Qy 4261 aaaaagccacattctgctctgggaaaaaataaaaaa 4301

Db 4261 AAAATGCCAATTTCTGTGCTCTGGGAAAAAATAAAAAA 4301

RESULT 3

HUMETR HUMETR 4286 bp mRNA linear PRI 23-JUN-1999

LOCUS Human mRNA for endothelin receptor (ETR), complete cds.

DEFINITION D90402

ACCESSION D90402.1 GI:219651

VERSION ETR; endothelin receptor.

KEYWORDS ETR; endothelin receptor.

SOURCE Human cDNA to mRNA.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 4286)

AUTHORS Ogawa,Y., Nakao,K., Arai,H., Nakagawa,O., Hosoda,K., Suga,S., Nakanishi,S. and Imura,H.

TITLE Molecular cloning of a non-isopeptide-selective human endothelin receptor

JOURNAL Biochem. Biophys. Res. Commun. 178 (1), 248-255 (1991)

MEDLINE 91298956

COMMENT Data kindly submitted in computer readable form by: Kazuwa Nakao

Second Division, Department of Medicine

Kyoto University School of Medicine

54 Shogoin Kawahara-cho

Sakyo-ku, Kyoto 606

Japan

Phone: 075-751-3170

Fax: 075-771-9452.

Location/Qualifiers

source

/organism="Homo sapiens"

/db_xref="taxon:9606"

238..1566

/codon_start=1

/product="endothelin receptor"

/protein_id="BAA14398.1"

/translation="MQPPPSLCGRALVALVACGLSRIMGEERGPDPDRATPLLOTAE

IMTPPTKTLWPKGSNLSARSLAPAEVPGKDRDTAGSPRTISPPPCOGPIEIKETFKY

INTVVCLVFLVGIIGNSTLLRIIYKNCMRNGPNILLASLALGDLHLHVIDIPINVY

KLLADWFFGAEMKALVFIQKASVGTIVLSICALSIDIRAVASWSRIKGIIVPKWT

AVEIIVWVSIVLVAPEAIGDIITMDYKSYLRICLLHPVOKFAFMQFYTKADWM

LFSFVCLPLAITAFYTMTCMELRKSGMOIALNDHLKORREVAKTFCVLVLVFLAL

CWLPLHLRLKLTLYNONDPNRCLESLFLVLDVYGINMASLNSINPIALVLSKR

FKNCFKSCLCWCQSQFEKQSLKFKANDHGYDNFRSSNKYS"

4286

polya_site

BASE COUNT 1327 a 828 c 817 g 1314 t

ORIGIN

Query Match

Best Local Similarity 99.7%; Score 4286; DB 9; Length 4286;

Matches 4286; Conservative 100.0%; Pred. No. 0;

Mismatches 0; Indels 0; Gaps 0;

Qy 1 gagacattcctgggggactctggccagcccgagcaacgctggatcctgagacactccc 60

Db 1 GAGACATTCGGTGGGGACTCTGCCACGCCGAGCAACGTGGATCCTGAGACACTCC 60

Qy 61 agtgaggcaattgccccgggtggagcgttgcagagcagtggtgtggcagggccccgtg 120

Db 61 AGGTAGGCATTTGCCCGGTGGAGCGCTTGCAGACGAGTGTGTGCAGGCCCCGTGG 120

Qy 121 aggtacacacagtgctgaacactgggaagaaactggtacttgagatctggacatctga 180

Db 121 AGGATCAACACAGTGGCTGAACACTGGGAAGAACTGGTACTTGGAGTCTGGACATCTGA 180

Qy 181 aacttggctctgaaactcggagcggccacggagccttcttgagcaggtagcagcatg 240

Db 181 AACTTGGCTCTGAACACTGGGAGCGGCCACCGGACGCCCTTCTGGACGAGTAGCACCATG 240

Qy 241 cagccgctccaagtctgtgcygagcgcctgtgtgctggttcttcttctgctgagcctg 300

Db 241 CAGCGCCTCCAAAGTCTGTGCGAGCGCGCCTGGTTGCGCTGGTCTTCTTGGCTGGCGCTG 300

Qy 301 tcgagatctgggag 360

Db 301 TCGCGGATCTGGGAG 360

Qy 361 accgagagataatgacccacccactaagacaccttatggcccaagggttcccaacgcaat 420

Db 361 ACCGACAGATAATGACGCCACCCACTTAAGACCTTATGCCCCCAAGGTTTCCAAACGCCAT 420

Qy 421 ctgctgagcgttctgtgacacctgagagagagagagagagagagagagagagagagagag 480

Db 421 CTGCGCGGCTGTTGGCACCTGCGGAGGTGGCTTAAAGGAGAGACAGACGCGAGGATCTCG 480

Qy 481 ccagcaccatctccctcccccgtgccaaggaccatcgaatcaagagagagagagagagagag 540

Db 481 CCAGCACCATCTCCCTCCCGGTGCGCAAGGAGGCCATCGAGATCAAGGAGAGATTTCAA 540

Qy 541 tacatcaacacggttctgctgcttctgctgcttctgctgctgctgctgctgctgctgctg 600

Db 541 TACATCAACAGGTTGTTGCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600

Qy 601 ctctgagaattatctacaagaacaagtgcagcgaaacggtcccaatatcttgcgcgc 660

Db 601 CTCTGAGAAATTATCTACAAGAACAAAGTGCATGCCAAACGGTCCCAATATCTTGTATCGCC 660

QY 661 agcttgctctggagaccctgctgcacatcgctcattgacatccctatcaatgtctacaag 720
DB 661 AGCTTGCTCTGGGAGACCTGCTGCACATCGTCAATTGACATCCCTATCAATATGCTACAAG 720
QY 721 ctctggcagagacgtggccatttgagctgagatggttaagctggtgcctttccatacag 780
DB 721 CTCTGGCAGGAGACTGGCCATTGGAGCTGAGATGTTAAGCTGGTGGCTTCATACAG 780
QY 781 aaagctcccgggaaacactgctgagtgctctatgctctgagtagttagacatatcga 840
DB 781 AAAGCCTCCGTGGGAATCACTGCTGAGTCTATGCTCTGAGTATTGACAGATATCGA 840
QY 841 gctgtgtctctctggagtagaataaaaggaaatgggggttccaaaatggagcagtagaa 900
DB 841 GCTGTGCTTCTTGGAGTAGAATTAAGGAATGGGGTTCCTCAAAATGGACAGCAGTAGAA 900
QY 901 attggtttgattgggtggctctctggtttctggtctgctgctgctgaagccataggtttgat 960
DB 901 ATTGTTTGAATTTGGGTGGTCTCTGTGGTTCTGGCTGTCCCTGAAGCCATAGGTTTGGAT 960
QY 961 ataattacgactgactacaagaaggtattctgcgaatctgcttgcctcccccgttcag 1020
DB 961 ATAATTACGATGGACTACAAAGGAAGTTATCTGCGAATCTGCTGTTCATCCCGTTCAG 1020
QY 1021 aagacagcttctcatgcagttttacaagacagcaaaagagattggtgctgttccagtttctat 1080
DB 1021 AAGACAGCTTTCATGCAAGTTTAAAGACAGCAAAAGATTGGTGGCTGTTCAGTTTCTAT 1080
QY 1081 tctgcttgccattggccatcactgcattttttatcacataatgacctgtgaaatttg 1140
DB 1081 TTCTGCTTGCCATTGGCCATCACTGCAATTTTTTATACATTAATGACCTGTGAAATGTTG 1140
QY 1141 agaagaaagtgatgcagattgctttaaagatcacctaaagcagagacggaagt 1200
DB 1141 AGAAGAAAGTGGCATGCAAGTTGCTTTAAATCATCACCTAAAGCAGAGACGGGAAGTG 1200
QY 1201 gcaaaaccgtctttgctggctctgtcttggcctctgctggtcttccccttcaacctc 1260
DB 1201 GCCAAAACCGCTTTTGGCTGGCTGCTTTTGGCCCTCTGCTGGCTTCCCTTTCACCTC 1260
QY 1261 agcaggattctgaagctcacctctttaaatacagaatgattcccaatagatgtgaactttg 1320
DB 1261 AGCAGGATTCTGAAGCTCACTCTTTATATCAGAAATGATCCCAATAGATGTGAACTTTG 1320
QY 1321 agctttctgtgtattggactatttggtatcaacatggcttccactgaattcctgcatt 1380
DB 1321 AGCTTTCTGTTGGTATTTGGACTATATTGGTATCAACATGGCTTCACATGAATTCCTGCATT 1380
QY 1381 aacccaattgctctgtatttggtagcaaaagattcaaaaactgctttaagtcattgctta 1440
DB 1381 AACCCAAATGCTCTGATTTGGTGAGCAAAAGATTCAAAAACCTGCTTTAAAGTCATGCTTA 1440
QY 1441 tgcctgctggcagtcatttgaagaaaaaacagctccttgagaaaaagcagctcgtgctta 1500
DB 1441 TGCTGCTGGTGCCAGTCATTTTGAAGAAAAACAGTCCCTTGGAGGAAAAAGCAGTCTGCTTA 1500
QY 1501 aagttcaaaagctaatgatcacgatatgacaaacttccgttccagtaataaatacagctca 1560
DB 1501 AAGTTCAAAGCTAATGATCAGGATATGCAAACTTCGTTCCAGTAATAAATACAGCTCA 1560
QY 1561 tcttgaagaagaactattcaactgtatttcaatttcttatttatttggaccgaagtcat 1620
DB 1561 TCTTGAAGAAGAACTATTCACTGTATTTCATTTCTTTTATTTATTTGGACCCGAAGTCAATA 1620
QY 1621 aacaaaatgaacatttggcaaaacaaaacaaataaactatgtatttgcacagcacatat 1680
DB 1621 AACAAAATGAACATTTGGCAAAACAAAACAAAACATATGATTTTGCACAGCACACATAT 1680
QY 1681 taataattaaagtgaattattttaaactcacagctcacatatgacattttatgagctgt 1740
DB 1681 TAAATATTAAAGTGAATATTATTAACTACAGCTACATATGACATTTATGAGCTGT 1740

QY 1741 ttacggcatggaaagaaaaacagtgggaaataagaaagccctcgtcgtgaaagcaactta 1800
DB 1741 TTACGGCATGGAAGAAAAATCAGTGGGAATTAAGAAAGCCCTCGTCTGGAAGCACTTAAT 1800
QY 1801 tttttacagttagcacttcaacatagctcttaacaacttccagatattcacacaacact 1860
DB 1801 TTTTTCAGTTAGCACTTCAACATAGCTCTTAACAACTTCAGGATATTACACAACACT 1860
QY 1861 taggcttaaaaaagcgtcactcagaatttctattcttctttaaagaagattttttta 1920
DB 1861 TAGGCTTAAAAATGAGCTCACTCAGATTCTATTCTTTCTAAAAACAGATTATTTTATA 1920
QY 1921 aatcaatgggactctgatataaaggagaataagtcactgtataaacagaacttttaaatg 1980
DB 1921 AATCAATGGGACTCTGATATAAAGGAAGAATAAGTCACTGTAAACAGAACTTTTAAATG 1980
QY 1981 aagcttaaaactcactcaatttaaaatttttaaaactttaaacaacttttcaataatat 2040
DB 1981 AAGCTTAAATTTACTCAATTTTAAATTTTAAATCTTTTAAACAACTTTTCAATTAATAT 2040
QY 2041 tatcacactattatcagattgtaattagatgcaaatgagagagcagttttagttgtgcat 2100
DB 2041 TATCACACTATTATCAGATTGTAATTAGATGCAAAATCAGAGAGCAGTTTAGTTCTGCAT 2100
QY 2101 ttttcggacactggaaacatttaaatgatcaggagggagtaacagaagaagcagggtgt 2160
DB 2101 TTTTCGGACACTGGAAACATTTAAATGATCAGGAGGGAGTAACAGAAGAGCAAGGCTGT 2160
QY 2161 ttttgaaaactcattcacacttctcactagaagccaaactcagcattctgcaatatgtaac 2220
DB 2161 TTTTGAATTCATTACACTTTTCACTAGAAGCCCAAACTCTCAGCATTCTGCATATGTAAAC 2220
QY 2221 caacatgtcacaaacaaagcagcatgtaacagactggcacatgtgcccagctgaaatttaaaa 2280
DB 2221 CAACATGTCAACAAACAGCAGCATGTAAACAGACTGGCACATGTGCCAGCTGAATTTAAAA 2280
QY 2281 tataactcttttaaaagaaaaattattacatcctttacattccttaagtaagatcaaacctca 2340
DB 2281 TATAACTTTTTTAAAAAGAAATTTATTACATCTTTTACATTTCACTTGAAGATCAAACTCA 2340
QY 2341 caaagaaataagaaattgttgaaaggctatcccaaaagacttttttgaactgtcattca 2400
DB 2341 CAAAGAAATAGAATCTTTGAAGGCTATCCCAAAAGACTTTTTCATCTGTCAATCA 2400
QY 2401 cataccctggaagacaataactatctacaattttttcaggatattataaaatctctctttt 2460
DB 2401 CATACCCTGGAAGACAATACTACTACTACAATTTTTCAGGATTTATTAATACTTCTCTTTT 2460
QY 2461 tcaactatcgtagctttaaactctgtttggttttgcctctgtaataacttaccctacataca 2520
DB 2461 TCACATATCGTAGCTTAAACTCTGTTGGTTTGTCTATCTGTAATACTTTACCTACATACA 2520
QY 2521 ctgcatgtagatgattaaatgaggcagccctgctgctatagcttccacatggagagat 2580
DB 2521 CTGCATGTAGTANTTAATGAGGCGAGCCCTGCTGCTCATAGCTTTTACGATGGAGAGAT 2580
QY 2581 gccagtgacctcataaataaagactgtgaactgctgctgctgctgctgctgctgctgctgctg 2640
DB 2581 GCCAGTGACCTCATATAATAAGACTGTGAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2640
QY 2641 gcaggtagaccctctctcccccattgctgtggttaaaatggtttctcagcagcagcagcagc 2700
DB 2641 GCAGGTAGCACCTCTCTCACCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2700
QY 2701 gctatagttaaaataactattttcaaaatcatcacagatttagtaccttaaacagctcctg 2760
DB 2701 GCTATAGTTAAAAATACATTATTTTCAAAATCATACAGATTTAGTACATTTACAGCTACCTG 2760
QY 2761 taaagcttattactaaattttttgtattttttgtaaatagccaatagaagaattttgtctg 2820
DB 2761 TAAAGCTTATTACTAAATTTTGTATTATTTTGTAAATAGCCAATAGCAAAAGTTTGTCTG 2820
QY 2821 acatggtgctttttcttctcatctcagagggcaaaactgctttttgagaccgtgaagaacctctt 2880

```
Db 2821 ACATGGTGTCTTCTTCTCATCTAGAGCAAACTGCTTTTGAGACCGTAAGAACCTCTT 2880
Qy 2881 agctttgtcgttcctgcctaaatttttatctcttaagcaaaagtccttagatagctt 2940
Db 2881 AGCTTTGTGCGTTCCTGCTAAATTTTATATCTTCTTAAGCAAAAGTCCTTAGATAGCTT 2940
Qy 2941 gggatgagatgtgtgaaagtatgtacaagagaaacggaagagagaaatgagggtg 3000
Db 2941 GGGATGAGATGTGTGTGAAGATATGTACAAGAGAAACGGAAGAGAGAAATGAGGTG 3000
Qy 3001 ggggtggaggaaacccatgggacagattcccatcttagcctaagcttcgctcattgctt 3060
Db 3001 GGGTGGAGGAAACCCATGGGACAGATTCCCATCTTAGCCTTAAGCTTCGTCATTGGCT 3060
Qy 3061 ggcacatcaatgcgaaggctcctgattttgtccagcaaaacacagtgcaatgttctca 3120
Db 3061 GGTCAATCAATGCAAAAGGCTCCTGATTTGTTCAGCAAAACACAGTGCATGTCTCA 3120
Qy 3121 gagtgaacttcgaataaattgggcccgaagagctttaactcgtgtcttaaaatatgcccaa 3180
Db 3121 GAGTGACCTTCGAATAAATTTGGGCCCAAGAGCTTTAACTCGGCTCTTAAATATGCCCAA 3180
Qy 3181 attttacttgttttctttaaagctggcgccacatgttggaataagctagtaagt 3240
Db 3181 ATTTTACTTGTCTTTCTTTAATAGGCTGGGCCACATGTTGGAATAAGCTAGTAATG 3240
Qy 3241 ttgtttctgtcaatattgaatgtgattgacagtaaaccaaaaccccaaatgtggcca 3300
Db 3241 TTGTTTTCTGTCAATATTGAATGTGATGTACAGTAACCAAAACCCCAACATGTGGCCA 3300
Qy 3301 gaaagaaagcaataaataattcaatccacacacatggtattctattataaacaccc 3360
Db 3301 GAAAGAAAGCAATAATAATTAATTCACACACCATATGGATTCATTATTAATCAACC 3360
Qy 3361 acaacttgttctttaaatttcaatcccaatcacttttccagggcctgtatcatagaagt 3420
Db 3361 ACAACTGTCTTAAATTTCAATCCCAATCCTTTTTCAGAGGCGCTGTATCATAGAAGT 3420
Qy 3421 cattttagactcctcaatttaaattaaatttgaatcactaataattttcacagttattaa 3480
Db 3421 CATTTAGACTCTCAATTTTAAATTAATTTTGAATCACTATATTTTACAGTTATTAA 3480
Qy 3481 tataattaaattctattaaatttagattattttttattaccatgactgaattttaca 3540
Db 3481 TATATTAATTTCTATTAAATTTTAGATTATTTTATTATACCATGACTGAATTTTACA 3540
Qy 3541 tcctgaacccttccctctccatgctcagatcatcgtctctcctaattattcttgcacaaatt 3600
Db 3541 TCCTGATACCTTTCCTTCTCCATGTCAGTATCATGTTCTCTAATTAATTCGCAAAATTT 3600
Qy 3601 tgaactacacacaaagacatactgtcattatttataataaaattgcattcagtgctt 3660
Db 3601 TGAACACTACACAAAAAGCATACTTGCATATTATTATAATAAATTCATTCAGTCAGTCTT 3660
Qy 3661 tttaaaaaaattgttgattcgaacttcaactactgataagtaagaacaaattataat 3720
Db 3661 TTTAAAAAAATGTTGATTCAAAACTTTAACACTGATGATAAGTAAGAAACAAATTAAT 3720
Qy 3721 ttctttacactcgaacaaacagatagaaaagtgctatcgttcaactcgaacaaatgt 3780
Db 3721 TTTCTTACATACNCAAAACCAAGATAGAAAAGGTGCTATCGTTCACATTTCAAAACATGT 3780
Qy 3781 ttcttagtattaaggactttaatatagcaacagacaaaaattattgtaacatgagttta 3840
Db 3781 TTCCTAGTATTAGGACTTTAATATAGCACACACAAAAATATTATTGTAACATGGATGTTA 3840
Qy 3841 cagctcaagatttataaagattttaacctatttctccctattattatccactgctaatt 3900
Db 3841 CAGCTCAAAAGATTATAAAGATTATAACCTATTTTCTCCCTTATTATACCATGCTAAT 3900
Qy 3901 gtagatgtattgttcaaacaccttttagtattgtagcttacaatggtgcaaaaggaataca 3960
```

```
Db 3901 GTGGATGTATGTTCAAAACACCTTTTAGTATTATAGCTTACATATGCCCAAGCAATACA 3960
Qy 3961 gtttatagcaaaacatgggtatgctgtagctaaactttkataaaagtgtataatacaatgt 4020
Db 3961 GTTTATAGCAAAAACATGGGTATGCTGTAGCTAACTTTTATAAAAGCTAATAACAATGT 4020
Qy 4021 aaaaaattatatctcgtggagagatttttggttgctcctaaagtggctatagttactgattt 4080
Db 4021 AAAAATATATATCTATCTGGGAGGATTTTGGTTGCCCTAAAGTGGCTATAGTTACTGATTT 4080
Qy 4081 ttattatgtaagcaaaacccaataaaaaatttaagtttttttaacaactaccctatttttc 4140
Db 4081 TTTATTATGTAAGCAAAAACCAATAAAAAATTAAGTTTAAAACTACCTATTATTTTC 4140
Qy 4141 actgtacacacactaattcatttaataactaattgattgttttaaaagaaataataatgtga 4200
Db 4141 ACTGTACAGACACTAATTAATAACTAATAATTAATTAATTAATTAATTAATTAATTAATTA 4200
Qy 4201 caagtggacattatttattgtaataatacaattatcaagcaagatgaaagtatttcaatt 4260
Db 4201 CAAGTGGACATATTATGTTAAATATACAAATTAATCAAGCAAGTATGAAGTTATTCANTT 4260
Qy 4261 aaaaagccacattctctgtctctggg 4286
Db 4261 AAAATGCCACATTTCTGCTCTGCG 4286
```

RESULT 4

```
S57283 LOCUS S57283 4286 bp mRNA linear PRI 28-JUN-1993
DEFINITION endothelin ET-B receptor [human, mRNA, 4286 nt].
ACCESSION S57283
VERSION S57283.1 GI:298321
KEYWORDS human.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4286)
AUTHORS Arai, H., Nakao, K., Hosoda, K., Ogawa, Y., Nakagawa, O., Komatsu, Y. and
Imura, H.
TITLE Molecular cloning of human endothelin receptors and their
expression in vascular endothelial cells and smooth muscle cells
Jpn Circ J 56 Suppl 5, 1303-1307 (1992)
MEDLINE 93180293
PUBMED 1291713
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 128424] from the original journal article.
This sequence comes from Fig. 5.
```

FEATURES

```
source Location/Qualifiers
1..4286
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..4286
gene
1..4286
CDS
338..1566
/gene="endothelin ET-B receptor, ET-BR"
238..1566
/gene="endothelin ET-B receptor, ET-BR"
/codon_start=1
/note="This sequence comes from Fig. 5; ET-BR"
/product="endothelin ET-B receptor"
/protein_id="AAB25531.1"
/db_xref="GI:298322"
/translation="MQPPSLCGRLVALVACLGRINGEERGFPPDPTPLQTAE
IMPPTKTLMPKGSNLSARSLAPVKGDRTAGSPRTISPPCGGPIKEKTFY
INTVSVCLFVLGIIGNLLRIYKNKMRNGNLIASLADGLHVIDIPINV
KLIAEDWPFGAEMCKLPFTOKASVGTIVLSICALSIDRYRAVASWSRIKIGVPKWT
AVEIVLWVSVVLAVPAIGFDIIIMDYKGSYLRICLLHPVQKTFMQFTKAKQMW
LRSFYCLPLAITAFYTLTCEMLRKKSQNALNDHLKORREKVAKTFCVLVLFAL
CMLPFLSLRLIKLTLYNQNDPNRCESLSLFLVDYIGINNASLNSCLNPALVLSKR
FNKCFKSCLCMCQCFEKEQSLKFKANDHGYDNFRSSNKYS"
BASE COUNT 1327 a 829 c 816 g 1314 t
ORIGIN
```

Query Match					
Best Local Similarity 99.6%; Score 4284.4; DB 9; Length 4286;					
Matches 4285; Conservative 0; Mismatches 1; Indels 0; Gaps					
Qy	1	gagacattccggtgggaactctgcacgcgcgcagcaggatcggtgatccctgagagcactccc	60		
Dy	1				
Dy	1	GAGACATTCCGGTGGGGACTGTGCCAGCCCGACAAGTGGATCTTGAGAGACTCCC	60		
Qy	61	aggtaggcatctgcccggtgggaaccttgcagagcagtgtgtggaagccccccgttg	120		
Dy	61	AGGTAGGCATTTCCCGTGGACGCCCTTGCCAGAGCAGTGTGTGGCAGGCCCCCGGTG	120		
Qy	121	aggatacaacacagtggctgaaacctggagaagaaactggtaacttgagtcctggacaatctga	180		
Dy	121	AGGATCAACACAGTGGCTGAACACTGGGAAGAACTGGTACTTGGAGTCTGGACATCTGA	180		
Qy	181	aacttgacctaaactcggagcgcgcacgcgcacgcctctggagcaggttagcacatg	240		
Dy	181	AAC TTGGCTCTGA AACTGGCGAGCGGCCACC GGACGCC TTTCTGGAGCACAGTAGCAGCATG	240		
Qy	241	cagcgctctccaagtctgtgcggaacgcgccttggctggcttgcctggcgccgtg	300		
Dy	241	CAGCGCCCTCCAAGTCTGTGCGACGCGCCCTGGTGGCTGTGCTTTCCTGCGCGCTG	300		
Qy	301	tgcggactcgtgggagagagagagagctccgcgcctgacagggccactcgccttttgaa	360		
Dy	301	TGCGGATCTGGGAGAGAGAGAGAGCGTTC CGCTGACAGGGCCACTCCGCTTTTGGAA	360		
Qy	361	accgagagataatgacgcacccactaagaccttatggcccaagggttcacaagccagt	420		
Dy	361	ACCGCAGAGTAATGACGCCACC ACTAAGACTTATGGCCCAAGGGTTC CAAAGCCAGT	420		
Qy	421	ctggcgcgctcttggcaactcgtgaggtgtgcctaagagacagcagcgaggaatctcgc	480		
Dy	421	CTGGCGCGTCTGTTGSCACTCTCGGAGGTGCCTAAGGAGACAGGACGGCAGGATCTCG	480		
Qy	481	ccaagcaactctcccctcccgcgtgcaagagaccactagatatcaagagagacttcaaa	540		
Dy	481	CCAGGCACCATCTCCCTCCCCCGTCCCAAGGACCCTATCGAGATCAAGSAGACTTCAA	540		
Qy	541	tacatcaacacggttgtctcgccttggctgtgtcgtcgtgggagatcatcgggaactccaca	600		
Dy	541	TACATCAACACGGTTGTGTCGTGCTTGTGCTGTGGGATCATCGGGAAC TCCACA	600		
Qy	601	ctcttgagaa ttatctacaagaacagtgcatagcgaacggtcccaaatcttgcctgcgc	660		
Dy	601	CTTCTGAGAA TTATCTAC AAGAACAAGTGCATGCCNAACGGTCCCAATATCTTGATGCC	660		
Qy	661	agcttgacctcgtggagacctgtgcacatcgtaattgacatccctatcaatgtctacaag	720		
Dy	661	AGCTTGGCTCTGGGAGACCTGTGACATCGTCAVTTGACATCCCTATCAATGCTCTACAAG	720		
Qy	721	ctgctggcagaggaactggccatttggagctgagatgtgtaaactggcttctcatacag	780		
Dy	721	CTGCTGGCAGAGACTTGGCCATTTGGAGCTGAGATGTGTAAGCTGGTGGCTTTCATACAG	780		
Qy	781	aaagcctccgctgggaatcactgtgctgagctctatgtgctcgtgattgacagatatcga	840		
Dy	781	AAAGCCTCCGCTGGCAATCACTGTGCTGAGTCTATGTGCTCTGAGTATTGACAGATATCGA	840		
Qy	841	gctgttgctctcttggagtagaattaagagaaattggggttccaaaattggacagcagtagaa	900		
Dy	841	GCTGTTGCTCTTTGGAGTAGAATTAAGAAATTTGGGTTCCAAAATGGACACAGCTAGAA	900		
Qy	901	attgttttgaattggggtgctctgtggttcttgctgtccctggaagcaataggttttgaat	960		
Dy	901	ATTGTTTGAATTTGGTGGTCTCTGTGTTTCTGGCTGTCCCTGAAGCCATAGGTTTGTAT	960		
Qy	961	ataattacgatgactacaaggaagtatctgcgaatctgtcttcatcccgcttcaag	1020		
Dy	961	ATAATTACGATGACATACAAGGAAGTTATCTGCGAATCTGTGCTTCAAGCCGTTTCAAG	1020		

Qy	1021	aagacagctttc	atgcagtgtt	ttaacaagacagcaaa	gaattggtggctgttc	aggttctcat	1080		
Db	1021	AAGACAGCTTT	CATGCAGTTTT	TACAAGACAGCAAA	GATTGTGGTGTTC	AGTTTCTCAT	1080		
Qy	1081	ttctgcttgcc	aattggccaactac	catttttttacactaat	gcccttgtaaa	tggttg	1140		
Db	1081	TTCTGCTTGCC	ATTGGCCAATC	ACTGCATTTTTT	TATACACTAAT	GACCTGTGAANTGTTG	1140		
Qy	1141	agaagaanaag	tggcatgcagatt	gctttaaatgatcacct	ataaacagagagcgga	gtg	1200		
Db	1141	AGAAGAANAAG	TGGCATGCAGATT	GCTTTAAATGATCACCT	ATAAACGACAGACGGGA	GTG	1200		
Qy	1201	gccaaaaacc	gtcttttggccgtg	ctctgtcttttggccctctg	ctgtggtctccccttc	acacctc	1260		
Db	1201	GCCAAAAACC	GTCTTTTGGCGTCTGT	CTGTGTCTTTGGCCCTCTG	CTGTGGCTTCCCCTTCA	CCCTC	1260		
Qy	1261	aqcagattcta	agaactcactctt	ataatcagaatgatcc	caatatagattga	acttttg	1320		
Db	1261	AGCAGATTCTA	GAAGCTCACTCTT	TATAATCAGAATGAT	CCCCAANTAGATGTGA	ACTTTTG	1320		
Qy	1321	agctttctgt	tggattggactata	tgtgtatacaactggtct	cactgaattcctgcatt		1380		
Db	1321	AGCTTTCCTGT	TGGTATTGGACTATATT	TGGTATCAACATGGCTT	CACTGAATTCCTGCATT		1380		
Qy	1381	aaccaaatgt	ctctgtatttgg	tgagcaaaagattc	aaaaactgtttaa	gtcatgctta	1440		
Db	1381	AACCCAAATTT	GCTGTGTATTGGTAG	CAAAAGATTCAA	AAACTGCTTTTAAGTCAT	GCTTA	1440		
Qy	1441	tgtctgtgtg	cagtcatttga	agaanaaacagctct	tggaggaagcagtc	gtgctta	1500		
Db	1441	TGCTGCTGTGG	CAGTCATTTGNAGA	AAAACAGTCTCTTGG	AGGAAAAGCAGTCTGCT	CTTA	1500		
Qy	1501	aggtcaaagc	taatgatcac	ygatgacaaactcc	gttccagtaataa	atacagctca	1560		
Db	1501	AGGTTCAAAGC	TAAATGATCAGCGAT	ATGACAACCTCCGTTC	CCAGTAATAAATAC	AGCTCA	1560		
Qy	1561	tcttgaaga	agaactatcact	gtgtatttcattctt	tatttatttggac	cggaagtcatta	1620		
Db	1561	TCTTGAAGAAGA	CAACTATTCACTGTATT	CTATTTCTTTATAT	TGGACCGAAGTCAT	TAA	1620		
Qy	1621	acaaaaatga	aacatttgc	aaaaaaca	aaaactatgtatt	tgcacagcacactat	1680		
Db	1621	ACAAAATGAAC	ATTGCCAAAACAACA	AAAAACTATGTAT	TGTCACAGGCACACTAT		1680		
Qy	1681	taaaatatta	agtgtaattatt	tttaacactcac	agctacatgacatt	ttagcgtgt	1740		
Db	1681	TAAATATTAA	TGTAAATTATTTT	TAACTCACAGCTAC	ATATGACATTTTAT	GAGCTGT	1740		
Qy	1741	ttacgcgat	tgaagaataatc	agtggaattaa	gaagcctcgtgtga	agacattaat	1800		
Db	1741	TTACGGCAT	TGGAAGAAAATC	AGTGGGAATTAA	GAAAGCCTCGCTG	TGAAGCACATTAA	1800		
Qy	1801	ttttacagt	tagcactca	aatagctct	taacaacttcc	aggatattcacaca	1860		
Db	1801	TTTTTACAGT	TAGCACTTCAACAT	ATAGCTTTAA	CAACTTCCAGGATATT	CACACAACACT	1860		
Qy	1861	taggcttaaaa	ttagctcactcaga	atttctattctt	cttaaaaaagagatt	tattttta	1920		
Db	1861	TAGGCTTAAAA	TGAGCTCACTCAGA	ATTCTATTCTTCTT	TCTANAAAGAGATTT	TTTTTA	1920		
Qy	1921	aatcaatg	ggactctgtata	aaaggaataatg	ctcgttaaacacaga	acttttaaatg	1980		
Db	1921	AATCAATGG	CACTCTCATATAA	AGGAAGAATAA	GATAGTCACTGTAA	AAACAGAACTTTTTAAATG	1980		
Qy	1981	aagcttaaat	tactcaattt	aaaaattt	aaaaatccctt	aaaaaacactttt	caattaat	2040	
Db	1981	AAGCTTAAAT	TACTCAATTTAA	AAATTTTAAAT	TCCCTTTAA	AAACCAACTTTTCA	ATTAATAT	2040	
Qy	2041	tatcacactat	tacagattgta	atttagatgc	aaatgagagacag	ctttagttgttg	1100	cat	
Db	2041	TATCACACTAT	TATACATTGTAA	TATAGATGCA	AAATGAGACGAGCT	TTTAGTTG	TGTCAT	2100	
Qy	2101	ttttcggac	acttgg	aaacattt	aaatgatc	aggagggag	taacagaa	gagcagctgt	2160

Db 2101 TTTTCGGACACATGGAACAACTTTAAATGATCAGAGGGAGTAACAGAAAGAGCAAGGCTGT 2160
QY 2161 ttttgaaatcatattacatttcaacttagaagcccaaacctcagcattctgcaataatgtaac 2220
Db 2161 TTTTGAAATCATTTACATTTTCACATAGAGCCCAACCTTCACATCTCTCAATATGTAAC 2220
QY 2221 caacatgtcaacaacaacagcagcatgtaacagactgcaacatgcccagctgcaatttaaaa 2280
Db 2221 CAACATGTCAACAACAACAGCAGCATGTAACAGACTGGCAGCATGTGCCAGCTGAATTTAAA 2280
QY 2281 tataactctttaaaaagaaatattacatctctttacattcagtttaagatcaaacctca 2340
Db 2281 TATATACTTTTAAAGAGAAATATTACATCTCTTTACATTCAGTTAAGATCAAAACCTCA 2340
QY 2341 caaagagaaatagaaatgttgaaggtatcccaaaagactttttgaaatctgtcatcca 2400
Db 2341 CAAGAGAAATAGAAATGTTGAAGGCTATCCCAAAAGACTTTTGAATCTGTCAATCA 2400
QY 2401 catacctgtgaagacaaatactatctacaattttttcaggattattaaaatctctctttt 2460
Db 2401 CATACCTGTGAAGACAACTATCTACAATTTTTCAGGATTATTAAAATCTCTCTTTT 2460
QY 2461 tcactatgtagctaaactgtttgtttgttgatctgttaactgttaactacatataca 2520
Db 2461 TCACATCTGTAGCTTAAACTCTGTGTTGTTGTCATCTGTAATACCTTACTACATACA 2520
QY 2521 ctgcatgtagatgattaaatgagcagcagccctatgctcatagctttacagatggagat 2580
Db 2521 CTGCATGTAGATGATTAATGAGGCGCGCCCTGTGCTCATAGCTTTACGATGGAGAT 2580
QY 2581 gccagtgacctcaataaagactgtgaactgcctggtgcagtgccacatgacaaaggg 2640
Db 2581 GCCAGTGACCTCAATAAAGACTGTGAACGTGCTGTGTCAGTGTCCACATGACAAAGGG 2640
QY 2641 qcaggtagcaccctctctcaacctgctgtgtgttaaaatggtttctagcatatgtataat 2700
Db 2641 GCAGGTAGCACCCTCTCTACCCATGTGTGGTTAAATAGGTTTCTAGCATATGTATAAT 2700
QY 2701 gctatagttaaaaactatttttcaaatcatcacagatttagcatatttaacagctacctg 2760
Db 2701 GCTATAGTTAAATACATATTTTCAAAATCATACAGATTAGTACATTTAAACAGTACCTG 2760
QY 2761 taaagcttatactaaatttttgtatattttttaaataagcacaatagaataaagttgcttg 2820
Db 2761 TAAAGCTTATTACTAAATTTTGTATTTTGTAAATAGCAATAGCAATAGAAAAGTTGCTTG 2820
QY 2821 acatggtctttcttctcatcagagcacaactgctttttgagaccgtgaagaacctctt 2880
Db 2821 ACATGGGCTTTTCTTTTCATCTAGAGGCAAACTGCTTTTGTAGACCGTAAGAAACCTCTT 2880
QY 2881 agctttgtgccttgcctcaatttttatcttcttaagcaaatgcttaggatagctt 2940
Db 2881 AGCTTTGTGGTTCCTGCCCTAAATTTTATATCTTAAAGCAAGTGCCTTAGGATAGCTT 2940
QY 2941 gggatgagatgtgtgaaagtagtacaagagaaaaacggaagagagagaaatgaggtg 3000
Db 2941 GGCATGAGATGTGTGAAGATATGTACAAGAGAAAACGGAGAGAGAGGAAATGAGGTG 3000
QY 3001 gggttggagaaacccatggggacagattcccaattcttagcctaacttgcctcattgcct 3060
Db 3001 GGGTTGGAGAAACCCATGGGACAGATTCCCATTTCTTAGCTTAACTGCTGCTATTGCCT 3060
QY 3061 cgtcacatcaatgcaaaaggtcctgattttgttccagcaaaaaacacagtgcaatgtctca 3120
Db 3061 CGTCACATCAATGCAAAAGTCTGTGATTTTGTCCAGCAAAACACAGTGCATGTTCTCA 3120
QY 3121 gagtgaacttccgaaataaattggcccaagagcttttaactcgtctttaaataatgccaa 3180
Db 3121 GAGTGACTTCGAATAAATTTGGGCCCAAGAGGCTTTTAACTCGGCTTTAAATATGCCAA 3180
QY 3181 atttttaactttgtttttctttaaaggctggggccacatgttggaaataagctagtaag 3240
|||||

Db 3181 ATTTTACTTTTGTCTTTTAAAGTGGGCCACATGTTGAAATAAGCTAGTAATG 3240
QY 3241 ttgttttctgtcaataattgaatgtggtacagtaaaacccaaacccaaatgtggcca 3300
Db 3241 TGTCTTCTGTCAATATTGAATGTGATGTACAGTAAACCCAAACCAATGTGGCCA 3300
QY 3301 gaaagaaagagcaataataattcaattcacacacacataggtattctattataaacccc 3360
Db 3301 GAAAGAAAGAGCAATAATAATTAAATTCACACACCATATGGATTCATTTATAAATCACCC 3360
QY 3361 acaactgttctttaatttcaatcccaatcactttttcagagcctgtgtatcataagaagt 3420
Db 3361 ACAAACTTGTCTTTAAATTTCAATCCCAATCACTTTTTCAGAGGCTGTTATCATAGAAGT 3420
QY 3421 caatttagactcgaattttaataatttaatttgaatcaactaatttttcacagtttataa 3480
Db 3421 CATTTTAGACTCTCAATTTTAAATTTTGAATCAGTAAATATTTTTCAGTATTATAA 3480
QY 3481 tataatttcttcttataatttttagattattttttaccatgtactgaatttttaca 3540
Db 3481 TATATTTAAATTTCTATTTAAATTTTAGATTTATTTTACCATGTACTGAATTTTACA 3540
QY 3541 tctgtatacccttctctctccatgtcagtatcatgttctcttaattatcttgcgaatttt 3600
Db 3541 TCCTGATACCCCTTCTCTCCATGTCAATATCATGTTCTCTAAATTTATCTTGCCAAATTT 3600
QY 3601 tgaactcacacacaaagcactcactgcaattttataataaaatgtcactcagtggtt 3660
Db 3601 TGAACACTACACACAAAGACATCTTGCATTTATTTATAATAAAATTTGCATTTCACTGCTT 3660
QY 3661 tttaaaaaaatggttgatcacaactttaacatactgataagtaagaacaattataat 3720
Db 3661 TTTAAAAAATGTTTGTATCAAACTTTAACATCTGATAGTAAGTAAGAAACAATTTAAT 3720
QY 3721 tctttacatactcaaaaaccaaagtagaaaaaggtgctatcgctcaacttcaaaaactgt 3780
Db 3721 TTTCTTACATCTCAAAAACCAAGATAGAAAAGGTGCTATCGTTCAACTTCAAAACATGT 3780
QY 3781 ttcctagattaaagaaactttaataagcaacagacaaaattattgttaacatggatgta 3840
Db 3781 TTCTTAGTATTAAAGACCTTAAATATAGCAACAGACAAAATTTATTTAAACATGGATGTTA 3840
QY 3841 cagctcaaaaagattttataaaagattttaacccatttttcccttatttccactgctgaat 3900
Db 3841 CAGCTCAAAAGATTTATAAAGATTTTAACTATTTTCTCCCTTATTTATCCACTGCTAAT 3900
QY 3901 gtgagtgatgttcaaacacaccttttagtattgtatgtagcttacaatgtggccaaaggaataca 3960
Db 3901 GTGGATGTATGTTCAACACACCTTTTAGTATTGATAGCTTACATATGGCCAAAGGAATACA 3960
QY 3961 gtttatagcaaaacatgggtatgctgttagctaaactttataaaagtgttaatacaaatgt 4020
Db 3961 GTTTATAGCAAAACATGGGTATGCTGTAGCTTAACTTTTATAAAGTGTATATATAACAATGT 4020
QY 4021 aaaaaattatatactctgggagatttttggctgaaggtgctatagttactgattt 4080
Db 4021 AAAAAATATATATCTGGGAGGATTTTGTGCTGAAGTGGCTATAGTTACTGATTT 4080
QY 4081 tttatattgaagcaaaccaataaaaaatttaagtttttttaaacactacacttttttc 4140
Db 4081 TTTTATTATGTAAAGCAAAACCAATAAAAAATTTAAGTTTTTTTAAACAACTTACCTTATTTTC 4140
QY 4141 actgtacagacactaatcttaataactaattatgattgttttaaaagaaataaaatgtga 4200
Db 4141 ACTGTACAGACACTTAATTCATTAATAACTAATTAATGTTGTTTTAAAAAGAAATATAAATGTGA 4200
QY 4201 caagtgcacattattatgttaataatacaattatcaagcaagatgaagtattcaatt 4260
Db 4201 CAAGTGGACATTTATTTATGTTAAATATACAAATTTATCAAGCAAGTATGAAGTTATTCAATT 4260
QY 4261 aaaaatgcccatttctggtctctggg 4286
Db 4261 AAAATGCCATTTCTGGTCTCTGGG 4286

RESULT 5
HUMHETBR7
LOCUS 2972 bp DNA linear PRI 14-APR-2000
DEFINITION Human gene for endothelin-B receptor (hET-BR), exon 7.
ACCESSION D13168
VERSION D13168.1 GI:285924
KEYWORDS G protein-coupled receptor; endothelin-B receptor; hET-BR.
SEGMENT 7 of 7
SOURCE Homo sapiens DNA, clone_lib:human genomic libraries.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 2972)
AUTHORS Arai, H., Nakao, K., Takaya, K., Hosoda, K., Ogawa, Y., Nakanishi, S. and Imura, H.
TITLE The human endothelin-B receptor gene. Structural organization and chromosomal assignment
JOURNAL J. Biol. Chem. 268 (5), 3463-3470 (1993)
MEDLINE 93155196
REFERENCE 2 (bases 1 to 2972)
AUTHORS Arai, H.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1992) Hiroshi Arai, Kyoto University School of Medicine, Second Division, Department of Medicine; 54 Shogoin, Kawahara-cho, Sakyo-ku, Kyoto, Kyoto 606, Japan (Tel:075-751-3170, Fax:075-771-9452)
COMMENT Submitted (02-SEP-1992) to DDBJ by:
Hiroshi Arai
Second Division, Dept. Of Medicine
Kyoto University School of Medicine
54 Shogoin Kawahara-cho
Sakyo-ku, Kyoto 606
Japan
Phone: 075-751-3170
Fax: 075-771-9452.
FEATURES Location/Qualifiers
1..2972
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="human genomic libraries"
join(D13162.1:1002..1742,D13163.1:11..123,
D13164.1:11..215,D13165.1:11..160,D13166.1:11..144,
D13167.1:11..119,11..2865)
/gene="hET-BR"
CDS
join(D13162.1:1260..1742,D13163.1:11..123,
D13164.1:11..215,D13165.1:11..160,D13166.1:11..144,
D13167.1:11..119,11..145)
/gene="hET-BR"
/codon_start=1
/product="human endothelin-B receptor"
/protein_id="BAA02445.1"
/db_xref="GI:285926"
/translation="MPPPSLCGRALVALVACLSRINGERGFPPDRATPLLTQAE
INTPRTKLWPGSNASLARSAPAEVPGKDRTAGPPRTISPPCCGPIEIKETKY
INTVSVCLFVFLIGITNLTIRIYKKNMRGNPILIASLALGDLHLHVIDIPINVY
KLLAEDWPEGAEMCKLVPTFKASVGTIVLSLCALSDIRAVASWSRINKGIGVKWT
AVEVILWVSVVLAVPEALMGDITIMDYKGVLRICLLHPVOKTAFMOFYKTAKDVK
LFSFYCLPLAIFATFYTLTMCMLRKKSGMQIALNDLHLKQREVAKTVFCLVLVAL
CWPLFLRLIKULTYNQNDPNCELLSFLVLVDYIGINNASLNSCNIPALYLVSKR
FNCFKSCLCMCQSFEEKQSLKEEKQSLCKFRKANDHYDNFRSSNKYSS"
11..2865
/gene="hET-BR"
/number=7
/evidence=experimental
/product="human endothelin-B receptor"
BASE COUNT 1018 a 499 c 465 g 990 t
ORIGIN Chromosome 13.

Query Match	66.4 %	Score 2857;	DB 9;	Length 2972;
Best Local Similarity	100.0 %	Pred. No. 0;		

	Matches	2857;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;								
Qy	1430	agtc	atgctt	atgctg	tgctg	tgccag	tcat	ttga	gaaaaa	caacag	tccctg	ggagaaa	gc	1489				
Db	9	AGT	CATGCTT	ATGCTGCTG	TGGTGCCAGT	CATT	TTGA	AAAAA	CAACAGT	CCCTG	TGCAGT	AAATG	68					
Qy	1490	agtc	gtctt	aaag	tccaa	agc	taata	gat	ca	cgg	atg	acac	ctcc	gtt	ccag	taata	1549	
Db	69	AGT	CGTCTT	AAAGTCT	CAAAGCT	AATGAT	CACGGAT	ATGACA	CACT	CCGT	TCCAGT	TAATA	128					
Qy	1550	aat	acag	ctcat	ctt	gaaa	aa	aa	aa	aa	aa	aa	aa	aa	aa	aa	1609	
Db	129	AAT	ACAGCT	CATCT	TGAA	GAAGA	AAAC	TATTC	AC	TGT	ATTT	TCAT	TTCTTT	TAT	TAT	TGACC	188	
Qy	1610	ga	gct	at	taaa	caaa	at	ga	aa	ca	ttt	g	ccaa	aa	caaa	aa	aa	1669
Db	189	GAAG	TCAT	TA	AAACAA	AAAT	GA	AA	CAAT	TG	CCAA	AA	CAAA	CAAA	AA	AA	AA	248
Qy	1670	cag	ca	ca	ca	ca	ca	ta	ta	ta	ta	ta	ta	ta	ta	ta	ta	1729
Db	249	CAG	CACAC	ACT	AT	TAAAA	TAT	AA	AGT	TAAT	TAT	TTT	TA	CACT	CAC	AGT	TACAT	308
Qy	1730	tt	at	gc	ctt	tt	ac	gc	at	tg	gaa	aaaa	aa	aa	aa	aa	aa	1789
Db	309	TTAT	GAGCT	GT	TAC	GGCA	TG	GA	AAAA	AAAT	T	CAGT	GG	AA	T	TAG	AA	368
Qy	1790	a	g	c	a	c	t	t	a	a	t	t	t	t	t	t	t	1849
Db	369	AAG	CAC	TTA	AT	TTTT	TAC	AGT	T	AG	CACT	T	CA	CA	T	AG	TAT	428
Qy	1850	c	a	c	a	a	c	a	c	t	t	a	a	a	a	a	a	1909
Db	429	CAC	AA	CA	CACT	T	AG	GT	T	AAAA	T	AG	CT	C	A	C	T	488
Qy	1910	a	t	t	a	t	t	t	t	a	a	a	a	a	a	a	a	1969
Db	489	AT	TTAT	TTTT	TA	AT	CA	T	GG	CA	CT	T	G	AT	TA	AA	AG	548
Qy	1970	a	c	t	t	t	a	a	a	t	t	t	t	t	t	t	t	2029
Db	549	ACT	TT	TA	AT	TGA	CT	T	AA	AT	T	T	AA	AT	T	T	AA	608
Qy	2030	t	c	a	a	t	a	t	a	t	a	t	a	t	a	t	a	2089
Db	609	TCA	AT	TA	AT	TAT	TAC	ACT	AT	TAT	TAC	AG	T	GT	T	AA	T	668
Qy	2090	a	g	t	t	g	c	a	t	t	t	t	t	t	t	t	t	2149
Db	669	AGT	T	GT	T	G	CA	T	TT	T	CG	CA	CT	T	G	GA	CA	728
Qy	2150	a	g	c	a	g	c	t	g	t	t	t	t	t	t	t	t	2209
Db	729	AG	CA	AGG	CT	G	T	T	T	T	G	A	A	A	A	A	A	788
Qy	2210	c	a	a	t	g	t	a	a	c	a	c	a	t	g	t	c	2269
Db	789	CA	AT	AT	G	T	AA	CA	CA	CA	CA	CA	CA	CA	CA	CA	CA	848
Qy	2270	t	g	a	a	t	t	t	t	t	t	t	t	t	t	t	t	2329
Db	849	TGA	AT	TT	AAAA	TAT	AA	T	ACT	TT	TT	AAAA	AG	AA	AT	TT	T	908
Qy	2330	a	t	c	a	a	c	t	c	a	a	g	a	a	a	a	a	2389
Db	909	AT	CA	AA	CT	CA	CA	AG	GA	AA	T	AG	A	T	G	T	T	968
Qy	2390	t	c	t	g	c	a	t	c	a	c	c	t	g	t	g	a	2449
Db	969	TCT	G	TC	AT	TC	AC											

QY	2510	a	ctacatacactgcgtagatgattaaatgagggcaggccctgctcatagctttac	256
Db	1089	ACTACATACACTGCATGCATGATGATTAATGAGGCGAGCCCTGTGCTCATAGCTTTTAC	1148	
QY	2570	gatgagaagatgccagtgacctataataaagactgtgaactgcctggcgagtgctccac	2629	
Db	1149	GATGGAGAGATGCCAGTGACCTCATTAATAAGACTGTGAACCTGCCGTGGTGCAGTGCCAC	1208	
QY	2630	atgcacaaagggcaggtagcacccctctccaccatgctgtggttaaaatggtttctagc	2689	
Db	1209	ATGACAAAGGGCAGGTAGCACCCCTCTACCCATGCTGTGGTGTAAATGGFTTTCTAGC	1268	
QY	2690	atatgataatgctagttaaataactatttttcaaaatcatcacagattagcacattta	2749	
Db	1269	ATATGTATAATGCTATAGTTAAATACTATTTTCAAAATCATACAGATTAGTACATTTA	1328	
QY	2750	acagctacctgtaaagcttattactaattttgtattattttgttaaatagcgaatagaa	2809	
Db	1329	ACAGCTACTCTGTAAAGCTTTATTACTAAATTTTGTATTATTTTGTATAAGCAATAGAA	1388	
QY	2810	aagtttgcctgacatgggtctttcttcatcatagaggcaaaacgcctttttgagaccgt	2869	
Db	1389	AAGTTTGCTTGACATGGTGTCTTCTCATCTAGAGGCAAAACGTCTTTTGTAGACCGT	1448	
QY	2870	aagaacctcttagctttgtgcttccctgcctaatttttatctctctaaagcaaatgacct	2929	
Db	1449	AAGAACCCTTTAGCTTTGTGGTTCCTGCCPAATTTTATATCTTCTAAGCAAAAGTGCCT	1508	
QY	2930	taggataagcttggagatgagatgtgtgtgaaagtatgtacaagagaaaaacgggaagagag	2989	
Db	1509	TAGGATAGCTTGGGATGAGATGTGTGTGAAAGTATGTACAGAGAAAAACGGGAAGAGAG	1568	
QY	2990	gaaatgagttgggtgtgagagaacccatgggacagattccattcttagcttaacggtt	3049	
Db	1569	GAATGAGTGGGGTGTGGAGGAACCCATGGGACAGATTCCTATCTTAGCTCAACGTT	1628	
QY	3050	cgctaatgctcgtccatcatcaatgcacaaaggtcctgattttgttcacgcaaaaacagtg	3109	
Db	1629	CGTATTTCCTCGTGCATCATCAATGCMAAGGCTCCTGATTTTGTCCACGCAAAACACAGT	1688	
QY	3110	caatgttctcagaagtgaacttcgaaaaataatgggcccagaagcttttaactcggtcttaa	3169	
Db	1689	CAATGTCTCAGAGTGACTTTCGAATAATAATTTGGGCCCAAGAGCTTTAACTCGGTCTTAA	1748	
QY	3170	aatatgcccaaaatcttcaactgttttcttctaataaggctgggccacatgttggaaata	3229	
Db	1749	AATATGCCCAAAATTTTACTTTGTTTTCTTTTAATAGGCTGGGCCACATGTTGGAAATA	1808	
QY	3230	agctagtaatgtgtttcttctgccaatatgaaatgtgatgtacagttaaaccaaacccaa	3289	
Db	1809	AGCTAGTAATGTGTTTTCTGTCAATATTGAATGTGATGGTACAGTAAACCAAAACCCAA	1868	
QY	3290	caatgtggccagaaaaagagagaataataattaattcacacaccatatgattctctattt	3349	
Db	1869	CAATGTGGCCAGAAAGAAAGAGCAATAATAATTAATTCACACACCATATGGATCTATT	1928	
QY	3350	ataaatcaccacaaaactgttcttaatttcaatttcaaccataactttttcagagagcctgtt	3409	
Db	1929	ATAAATCACCCACAACCTGTTCTTTAAATTTTCATCCCAATCACTTTTTCAGAGGCCCTGT	1988	
QY	3410	atcatagaagtcattttagactctcaattttaaattaattttgaaatcaactaataatttca	3469	
Db	1989	ATCATAGAAGTCATTTTAGACTCTCAATTTTAAATTAATTTTGAATCACHAATAATTTC	2048	
QY	3470	caagtttataataatttcaatttctattbaaattttagattatttttattaccatgtact	3529	
Db	2049	CAGTTTATTTAATATATTTAATTTCTATTATAATTTTAGATTATTTTATTAACCATGTACT	2108	
QY	3530	gaattttacatcctgatacccttctctccatgtccatgtcagatcatcgttctctaattatc	3589	
Db	2109	GAATTTTACATCCCTGATACCCCTTCCCTTCCTCATCTGAGTCAATGTCTCTCTTAATATC	2168	

QY	3590	ttgccaaattttgaaactacacacaaaagcatacttgctatttataataaaaattgca	3649
Db	2169	TTGCCAAATTTTGAAACTACACACAAAAGCATCTTGCAATTTTATATAAAATTGCA	2228
QY	3650	ttcagtggtgttttaaaaaaaatgtttgattcacaacaccttaacatactgataagtaa	3709
Db	2229	TTcAGTGGCTTTTAAAAAAATGTTTGATTCAAAACCTTTTAACATACTGATAAGTAAGAA	2288
QY	3710	acaattataattctttcacatactcaaaaaccagaatagaaaaagtgctatccttcaact	3769
Db	2289	ACAATTATAATTCTTTACATACTCAAAACCAAGATAGAAAAGGTGCTATCCTGTTCAACT	2348
QY	3770	tcaaacatgtttcctcagttattaaaggacctttaatagcaacagacacaaattattgttaa	3829
Db	2349	TCAAAACATGTTTCTCTAGTATTAAAGGACTTTAATATAGCAACAGACAAAATTATTGTTAA	2408
QY	3830	caatgatgttacagctcaaaaagatttataaaagattttaacctattttcctccttattat	3889
Db	2409	CATGGATGTTTACAGCTCAAAAGATTATATAAGATTTTAACCTATTTTCTCCTTATTAT	2468
QY	3890	ccactgctaaatgtggatgtatgtcacaacaccttttagttattgatagcttacatatggcc	3949
Db	2469	CCACTGCTTAATGTGATGTATGTTTCACACACCTTTTAGTATTGTATAGCTTACATATGGCC	2528
QY	3950	aaaggaatcacagttttatagcaaaaacatgggtatgctgtagctcaactttataaagtgtaa	4009
Db	2529	AAAGGAATACAGTTTATAGCAAAACATGGGTATGCTGTAGCTTAACTTATAAAGTGTAA	2588
QY	4010	tataacaatgtaaaaaatatatactctggaggattttttggttgcctaaagtggtcata	4069
Db	2589	TATAACAATGTAAAAAATTATATATCTGGGAGGATTTTGTGTTGCCTAAGCTGCTATA	2648
QY	4070	gttactgattttttattatgtaagcaaaaccaaataaaaatttaaagttttttaacaacta	4129
Db	2649	GTTACTGATTTTTTATTATGTATAGCAAAACCAATAAAAATTTAAGTTTTTTTAAACAACATA	2708
QY	4130	ccatttttccactgtacagacactaaattcatataactaattgatgtgttttaaagaaa	4189
Db	2709	CCTTATTTTTTTCACGTACAGACACTAATTCATTAAATGATTGTTTTTAAAGAAA	2768
QY	4190	tataaatgtcagaagtggaactattttatgtttaaataatacaattatcaagcaagtatgaa	4249
Db	2769	TATAAAATGTCAAAAGTGGACATTATTTATGTTTAAATATACAATTTATCAAGCAAGTATGAA	2828
QY	4250	gttattcaattaaatgccacatttctggtctctcgtgg	4286
Db	2829	GTTAATTCAAATTAATGCCACATTTCTGGTCTCTCGGG	2865

```

RESULT      6
AL139002/c
LOCUS
DEFINITION Human DNA sequence from clone Rpl1-318G21 on chromosome
              183337 bp DNA linear PRI-28-JAN-2001
ACCESSION  AL139002
VERSION    AL139002
KEYWORDS   13q22.2-31.1, complete sequence.
SOURCE     AL139002.18 GI:12597038
            human..
ORGANISM   Homo sapiens
            Eukaryota; Chordata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 183337)
            Wall,M.
            Direct Submission
            Submitted (28-JAN-2001) Sanger Centre, Hinxton, Cambridgeshire,
            CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
            requests: clonerequest@sanger.ac.uk
            On Jan 29, 2001 this sequence version replaced gi:12584355.
COMMENT

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with

only a small overlap as described above.
 This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated repeat sequence elements. Where the sequence is ambiguous, there is an annotation using the 'unsure' feature key. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; ENBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 13, constructed by the Sanger Centre Chromosome 13 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr13> RP11-318G21 is from the library RP11-11.2 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

FEATURES

Source

Location/Qualifiers
 1..183337
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="13"
 /map="q22.2-31.1"
 /clone="RP11-318G21"
 /clone_lib="RP11-11.2"
 3..454
 /note="LIMC5 repeat: matches 7127..7575 of consensus"
 repeat_region
 1216..1308
 /note="HAL1 repeat: matches 1475..1563 of consensus"
 repeat_region
 1309..1597
 /note="AluJb repeat: matches 1..297 of consensus"
 repeat_region
 1598..2044
 /note="HAL1 repeat: matches 1003..1475 of consensus"
 repeat_region
 2148..2276
 /note="L2 repeat: matches 2620..2749 of consensus"
 repeat_region
 2330..2378
 /note="L2 repeat: matches 2442..2492 of consensus"
 repeat_region
 3915..4224
 /note="AluY repeat: matches 1..306 of consensus"
 repeat_region
 4617..4750
 /note="67 copies 2 mer cc 61% conserved"
 repeat_region
 4648..4727
 /note="20 copies 4 mer cctt 78% conserved"
 repeat_region
 4729..4784
 /note="14 copies 4 mer toct 78% conserved"
 repeat_region
 5431..5736
 /note="AluX repeat: matches 1..305 of consensus"
 repeat_region
 11990..12273
 /note="AluX repeat: matches 9..292 of consensus"
 repeat_region
 12589..12809
 /note="MIR repeat: matches 7..234 of consensus"
 repeat_region
 13390..13519
 /note="L2 repeat: matches 2410..2548 of consensus"
 repeat_region
 14630..14978
 /note="THE1B repeat: matches 1..360 of consensus"
 repeat_region
 15092..15580
 /note="LIMB1 repeat: matches 5656..6116 of consensus"
 repeat_region
 15581..16095
 /note="L1P47 repeat: matches 5629..6143 of consensus"
 repeat_region
 16096..16549
 /note="LIMB1 repeat: matches 5188..5656 of consensus"
 repeat_region
 16731..16777
 /note="MIR repeat: matches 35..78 of consensus"
 repeat_region
 16778..17137
 /note="THE1C repeat: matches 1..371 of consensus"
 repeat_region
 17138..17273
 /note="MIR repeat: matches 78..226 of consensus"
 repeat_region
 17374..17484
 /note="MIR repeat: matches 26..158 of consensus"

repeat_region
 17485..17777
 /note="AluSc repeat: matches 1..290 of consensus"
 repeat_region
 17778..17815
 /note="MIR repeat: matches 158..191 of consensus"
 repeat_region
 19801..19048
 /note="34 copies 2 mer tt 66% conserved"
 repeat_region
 19447..19589
 /note="MIR repeat: matches 131..262 of consensus"
 repeat_region
 19843..20162
 /note="MER33 repeat: matches 1..324 of consensus"
 repeat_region
 20866..21198
 /note="MER44A repeat: matches 3..333 of consensus"
 repeat_region
 21742..21878
 /note="MIR repeat: matches 9..154 of consensus"
 repeat_region
 22214..22310
 /note="MIR repeat: matches 164..260 of consensus"
 repeat_region
 2321..22418
 /note="LIMB8 repeat: matches 6078..6171 of consensus"
 misc_feature
 22390..22715
 /note="Sequence from AC018674 sequenced by WUGSC."
 repeat_region
 22419..22730
 /note="AluY repeat: matches 1..311 of consensus"
 repeat_region
 22731..23714
 /note="LIMB8 repeat: matches 5130..6078 of consensus"
 repeat_region
 23715..24008
 /note="AluSg repeat: matches 1..294 of consensus"
 repeat_region
 24009..24264
 /note="LIMB8 repeat: matches 4884..5130 of consensus"
 repeat_region
 24265..24569
 /note="AluY repeat: matches 1..305 of consensus"
 repeat_region
 24570..25577
 /note="LIMB8 repeat: matches 3786..4884 of consensus"
 repeat_region
 25582..25635
 /note="27 copies 2 mer tt 70% conserved"
 repeat_region
 26221..26571
 /note="MSTA repeat: matches 1..347 of consensus"
 repeat_region
 26572..26750
 /note="AluY repeat: matches 129..307 of consensus"
 repeat_region
 26752..27066
 /note="AluY repeat: matches 1..311 of consensus"
 repeat_region
 27067..27134
 /note="MSTA repeat: matches 347..371 of consensus"
 repeat_region
 27963..28006
 /note="22 copies 2 mer tt 75% conserved"
 repeat_region
 28916..29282
 /note="MER39 repeat: matches 13..381 of consensus"
 repeat_region
 29282..29517
 /note="MER39B repeat: matches 327..579 of consensus"
 repeat_region
 30210..30526
 /note="AluJo repeat: matches 1..303 of consensus"
 repeat_region
 31423..31572
 /note="L1P13 repeat: matches 6005..6155 of consensus"
 repeat_region
 31587..31624
 /note="19 copies 2 mer tt 86% conserved"
 repeat_region
 32103..32181
 /note="ORSL repeat: matches 390..467 of consensus"
 repeat_region
 33878..34312
 /note="MER57A repeat: matches 1..433 of consensus"
 repeat_region
 36673..36768
 /note="L1R37A repeat: matches 81..172 of consensus"
 repeat_region
 36769..37066
 /note="AluSg repeat: matches 1..296 of consensus"
 repeat_region
 37067..37300
 /note="L1R37A repeat: matches 172..424 of consensus"
 repeat_region
 39470..39501
 /note="16 copies 2 mer tt 90% conserved"
 repeat_region
 41434..42607
 /note="L1M4 repeat: matches -258..888 of consensus"
 repeat_region
 42744..43220
 /note="L1M4 repeat: matches 1085..1580 of consensus"
 repeat_region
 43703..44007
 /note="AluJb repeat: matches 1..305 of consensus"
 repeat_region
 44019..44180

Db	2101	NNNNNTGTTGATTCAAACCTTTAAACATACTGATAAGTAAGAAACAATTAATAATTCCTTT	2160	
Qy	3727	acatactcaaaacccaagatagaaaaaggtgctatcgttcaacttcaaaacatgtttccta	3786	
Db	2161	ACATACTCAAAACCAAGATAGAAAAGGTCTATCGTTCAACTTCAAAACATGTTCCCTA	2220	
Qy	3787	gtattaagaccttaatatagacaagacaaaaattattgttaacatggattacagctc	3846	
Db	2221	GTATTAAGGACTTTAATATATAGCAACACACAAAATTTATGTTAAACATGGATGTTACAGCTC	2280	
Qy	3847	aaaagattataaaaagatttaacctattttcccttattatccactgctaatgtggat	3906	
Db	2281	AAAAGATTTATAAAGATTTTAACCTATTTCCTCTATTATCCACTGCTAATGTGGAT	2340	
Qy	3907	gtatgttcaaacacccttttagtattgatagcttacatatggccaaaggaatacagtttat	3966	
Db	2341	GTATGTTCAAAACACCTTTTATGATTGATAGCTTACATATGCGCAAGGAATACAGTTTAT	2400	
Qy	3967	agcaaaacatcgggtatgctgctagctaaactttataaagtgtaataatacaaatgtaaaaa	4026	
Db	2401	AGCAAAACATCGGTATGCTGTAGCTAACTTATATAAAGTGTATATAACAATGTAAAAA	2460	
Qy	4027	ttatatctcggagagatttttgggtgcctaaagtggtctatagttactgatttttatt	4086	
Db	2461	TTATATATCTGGGAGGATTTTGGTTGCTTAAAGTGGCTATAGTTACTGATTTTATT	2520	
Qy	4087	atgtaagcaaaacccaataaaaaattaaagtttttttaacaactacccttattttcacgtga	4146	
Db	2521	ATGTAACCAAAACCAATAAAAAATTTAAGTTTTTTAAACAACACTACCTTTATTTCACGTGA	2580	
Qy	4147	cagacactaattcaataataactaactgattgttttaaaagaaataataatgtgacaagtg	4206	
Db	2581	CAGACACTAATTCATTAAATACTAATTTGATTGTTTAAAGAAATATATAATGTGACAAAGT	2640	
Qy	4207	gacattatttatgttaaatatacaattatcaagcaagctatgaagttattcaattaaaatg	4266	
Db	2641	GACATTATTATGTTAAATATAAATAATATCAAGCAAGTATGAAGTTATTCAATTAATAATG	2700	
Qy	4267	ccacatttctggtctctggg 4286		
Db	2701	CCACATTTCTGCTCTGGG 2720		
RESULT	8			
AR165435				
LOCUS	AR165435	1873 bp	DNA	linear
DEFINITION	Sequence 13 from patent US 6280931.			PAT 17-OCT-2001
ACCESSION	AR165435			
VERSION	AR165435.1	GI:16240327		
KEYWORDS	Unknown.			
SOURCE	Unknown.			
ORGANISM	Unclassified.			
REFERENCE	1 (bases 1 to 1873)			
AUTHORS	Sakamoto,A. and Hanaoka,F.			
TITLE	Method for specifically amplifying a dystroglycan, .alpha.-sarcoglycan, or endothelin Breceptor cDNA of an extremely small			
JOURNAL	Patent: US 6280931-A 13 28-AUG-2001;			
FEATURES	Location/Qualifiers			
source	1. .1873			
BASE COUNT	490 a 434 c 438 g 511 t			
ORIGIN				
Query Match	39.3%	Score 1691.8;	DB 6;	Length 1873;
Best Local Similarity	99.6%;	Pred. No. 2.6e-292;		
Matches 1696;	Conservative	0; Mismatches	7; Indels	0; Gaps
0;				
Qy	178	tgaacctgtgctgaacctgaggagggccaccggagcgtctctgtgagcaggttagcagc	237	

Db	171	TGTCCTAGGCTCTGAAACTGCGAGCGGCCACCGNACGCCCTTCTGGAGCAGGTAGCAGC	230	
Qy	238	atcagccgctccaagtctgtggaacgccccctgggttgcgtggttcttgcctgcgc	297	
Db	231	ATCAGCGCGCTCCAAGCTCTGCGGACCGGCTGTTGCGTGTCTTGGCTGCGGC	290	
Qy	298	ctctcggaatctgggagagagagaggtctcccgctacagggccactcgcctttg	357	
Db	291	CTGTGCGGATCTGGGAGAGAGAGAGGCTTCCCGCTGACAGGGCCATCCGCTTTG	350	
Qy	358	caaacgcagagataatgagccaccactaaagaccttatggccaaaggttccaacgcc	417	
Db	351	CAAAACGCACAGATAATGACGCCACCTTAAGACCTTATGGCCCAAGGTTCCAACGCC	410	
Qy	418	agctcgccggtcggttggcaactgagaggtgcctaaagagagacagcaggaatct	477	
Db	411	AGTCTGGCGGGTCTGGCACCTGCGGAGGTGCTTAAAGGAGACAGAGCGGAGGATCT	470	
Qy	478	ccgccagcaccatctccctcccgctccaaaggaccatcgagatcaaggagacttc	537	
Db	471	CCGCCAGCACCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTTC	530	
Qy	538	aaatacatcaaacacggttgtctgctgtgttctgctggtggtggtatcgcggaactcc	597	
Db	531	AAATACATCAACACGGTTGTCTCTGCTTGTGTTCTGCTGGGATCATCGGGAACCTC	590	
Qy	598	acactctgagaattatctacaagaacaagtgcagtgcgaacggtcccaatatattgac	657	
Db	591	ACACTTCTGAGAATTTATCTACAAGACAAGTGCATGCGAAACGGTCCCAATATCTTGATC	650	
Qy	658	gccagcttgctctggagaccctgctcacatcgtcattcaatccatccatcaatgtctac	717	
Db	651	GCCAGCTTGCTCTGGGAGACCTGCTGCATCTGATTCATGACATCCCTATCAATGTCTAC	710	
Qy	718	aagctgctggcagagactggccatttggagctgagatgtgtgaagctgggtgcttccata	777	
Db	711	AGCTGCTGGCAGAGACTGGCCATTTGGAGCTGAGATGCTAGCTGGTGGCTTTCATA	770	
Qy	778	cagaagcctccgtgggaaactcactgctgagctctatgctctgagattgacagatat	837	
Db	771	CAGAAAGCCTCCGTGGGAATCACTGTGCTGAGTCTATGCTCTGAGTATTTGACAGATAT	830	
Qy	838	cgagctgttctcttgagtagaattaaaggaattgggttccaaatggacagcagta	897	
Db	831	CGAGCTGTCTCTTGGAGTAGAATTTAAAGGAATGGGGTTCCAAATGGACAGCAGTA	890	
Qy	898	gaaatttcttgattgggtggtctctggttctggtctgctccctgaagccatagtttt	957	
Db	891	GAATTTGTTTGTGTTGGTGGTCTCTGTGGTCTGCTCCCTGAAGCCATAGGTTTT	950	
Qy	958	gataaattacgatggactacaaagggaagtattatctgcgaatctgctgtcttccctgtt	1017	
Db	951	GATATAATTACGATGGACTACAAAGGAAGTTATCTGCGAATCTGCTTGTCTCATCCGCT	1010	
Qy	1018	cagaagacagcttcttcagcagttttacaagacagcaaaagattggtggctgttcagtttc	1077	
Db	1011	CACAAGACAGCTTTTCATGCGAGTTTTCACAAGACAGCAAAAGATTGGTGGCTATTAGTTTC	1070	
Qy	1078	tattctgctggcattggccatcactgcattttttatacactaaatgacctgtgaaatg	1137	
Db	1071	TATTTCTGCTTGGCATTGGCCATCAGTCGATTTTTTTATACACTAATGACCTGTGAAATG	1130	
Qy	1138	ttgagaagaaaagtgccatgcagattgcttttaaatgatcacctaaagcagagacggaaa	1197	
Db	1131	TTGAGAAAGAAAAGTGGCATGCGAGATTGCTTTAAATGATCACCTTAAAGCAGACAGCGGAA	1190	
Qy	1198	gtggcaaaaccgtcttttgcctggtctctgttcttccctctgctggttcccccttcac	1257	
Db	1191	GTGGCCAAAACCGCTTTTTCCTGGTCTCTGTCTTCTTCCCTCTGCTGGCTTCCCTTCAC	1250	
Qy	1258	ctcagcaggtattctgaagctcactctttataatcagatgatcccaatagatgtgaact	1317	
Db	1251	CTCAGCAGGATCTGAAGACTCAGTCTTTTAAATCAGAAATGATCCCATAGATGTGAACCT	1310	

```

QY 1318 ttgagctttctgttggttgactatattggtatcaacatgggtctcaactgaattcctgc 1377
Db 1311 TTGAGCTTCTGTGGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATCCTGC 1370
QY 1378 attaaacccaattgctctgtatttggtgagaaaagattcaaaaactgctttaagtcacgc 1437
Db 1371 ATTAACCCAATTGCTCTGTATTGGTGAGCAAAAAGATTCAAAAACACTGCTTTAAGTCATGC 1430
QY 1438 ttatgctgtggtgacagtcatttgaagaaaaaacagtccttgaggaagaaacagtcctgc 1497
Db 1431 TTATGCTGCTGGGCCAGTCACTTTGAAGAAAACAGTCCCTTGGAGAAAAGCAGCTCGTGC 1490
QY 1498 ttaaagtccaagtaattgatcacgggatgacaaacttccggttccagtaataaaatacagc 1557
Db 1491 TTAAGTTCAAGCTAATGATCAGGATATGACAACCTTCGTTCCAGTAATAATACAGC 1550
QY 1558 tcatottgaagaagaactattcaactgatttcaatttcttatttatttggaccggaagtcac 1617
Db 1551 TCATCTTTGAAGAAGAACTATTCACTGTATTTCATTTTCTTTATATTGGACCCGAAGTCAT 1610
QY 1618 taaacaaaatgaacatttgcacaaacaaacaaacaaacaaacaaacaaacaaacaaac 1677
Db 1611 TAAACAAATGAACATTTGCCAAACAAACAAACAAACAAACAAACAAACAAACAAAC 1670
QY 1678 tattaaaaattaaagttaatttttaacactcacagctcacatatacatttttatgagc 1737
Db 1671 TATTAAAAATTAAAGTGAATTATTTAACACTCACAGCTACATATGACATTTTATGAGC 1730
QY 1738 tgtttacgcatggaagaataacagtggaataaagaagcctcgtcgtgaagcaactt 1797
Db 1731 TGTTTACGGCATGGAAGAAAATCAGTGGGAATTAAGAAAGCCTCGTTCGTGAAAGCACTT 1790
QY 1798 aattttttagcttagcacttaacatagctttaaacaacttccaggatattcacacaac 1857
Db 1791 AATTTTTACAGTTACACTTCAACATAGCTCTTAACAACCTTCAGGATATTACACAAC 1850
QY 1858 acttaggcttaaaaatgagctca 1880
Db 1851 ACTTAGGCTTAAAAATGAGCTCA 1873

RESULT 9
E15242
LOCUS Human mRNA for endothelin B receptor, complete cds. 1873 bp DNA linear PAT 28-JUL-1999
DEFINITION
ACCESSION E15242
VERSION E15242.1 GI:5709925
KEYWORDS JP 1998057064-A/13.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1873)
AUTHORS Sakamoto, E. and Hanaoka, F.
TITLE SPECIFIC AMPLIFICATION OF MINOR GENE PRODUCT
JOURNAL Patent: JP 1998057064-A 13 03-MAR-1998;
RIKAGAKU KENKYUSHO
COMMENT OS Homo sapiens (human)
PN JP 1998057064-A/13
PD 03-MAR-1998
PF 16-AUG-1996 JP 1996216506
PI SAKAMOTO EIJI, HANAOKA FUMIO
PC C12N15/09,C07H21/02,C07H21/04//C12Q1/68;
CC strandedness: Double;
topology: Linear;
Key Location/Qualifiers
source 1..1873
/organism="Homo sapiens"
/tissue_type="peripheral blood" FT CDS
FT 231..1559
/product="endothelin B receptor".

```

```

FEATURES
source 1..1873
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 490 a 434 c 438 g 511 t
ORIGIN
Query Match 39.38; Score 1691.8; DB 6; Length 1873;
Best Local Similarity 99.68; Pred. No. 2.6e-292;
Matches 1696; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 178 tgaacttggctctgaaactcgagcgccacacggagcctctctggagcagtgtagcagc 237
Db 171 TGTCTCTAGGCTCTGAACACTGCGGAGCGGCCACCGGACGCCCTTCTGGAGCAGGTAGCAGC 230
QY 238 atgcagcgcgcctcaagtctgtgcgagcgccctggttgcgtggttcttcttgcctgcgc 297
Db 231 ATGCAGCGCGCTCCAAAGTCTGTGCGGACCGGCCCTGTTGCGCTGGTCTTGGCTTCGCGC 290
QY 298 ctgtcgcggactctgggagagagagaggtctccgcctcgacagggccactccgcgttttg 357
Db 291 CTGTCGGCGATCTCGGGAGAGAGAGAGGCTTCCCGCCTGACAGGGCCACTCCGCTTTTG 350
QY 358 caaacccgcagagataatgacgcccacactaaagaccttatggcccaagggttccaacgccc 417
Db 351 CAAACCCGACAGATATATGACGCCACCCACTTAAGACCTTATGCGCCCAAGGGTTCCAACGCC 410
QY 418 agtctggcgcggttgcttggcacctcgcgaggtgcttgaagagagacagacgagcaggtatct 477
Db 411 AGTCTGCGCGGTGCTTGGCACCTTGGGAGGTGCTAAAGGAGACAGGACGCGAGGATCT 470
QY 478 ccgcccagcaccatctccctccctcccgctgcaagagaccatcgagatcaagagacatttc 537
Db 471 CCGCCAGCCACCACTCTCCCTCCCGCTGCCAAGAGCCCAATCGAGATCAAGGAGACTTTC 530
QY 538 aaatacacaacacggtgtgctcctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 597
Db 531 AATATACATCAACAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
QY 598 acacttctgagaattatctacaagaacaagtgcagtgcgaaacggtcccaatatcttgatc 657
Db 591 ACACTTCTGAGAAATTATCTACAAGAACAAAGTGCATGCGAAACGGTCCCAATATCTTGATC 650
QY 658 gccagctgtgctctgggagacctgtgcacatgcattgaacatccctatacaatgtctac 717
Db 651 GCCAGCTTGGCTCTGGGAGACCTGCTGCATCTCATTTGACATCCCTATCAATGTCTTAC 710
QY 718 aagctgctggcagagactggccattggagctgagatgtgtgaagctggtgctcttcata 777
Db 711 AAGCTGCTGCAGAGGACTGGCCATTTGGAGCTGAGATGTGTAAAGCTGGTGCCCTTTTCATA 770
QY 778 cagaaagcctccgtgggaatcacctgtgctgagtgtatgtctgtgctgagtattgacagatat 837
Db 771 CAGAAAGCCCTCCGTGGGAATCACTGTGCTGAGTCTATGCTCTGAGTATTGACAGATAT 830
QY 838 cgagctgtgctctctggagtagaataaaggaaattgggggttccaaaatgagacagcagta 897
Db 831 CGAGCTGTTGCTCTTGGAGTAGAATTAAGGAATTTGGGGTTCCAAAATGACACAGCAGTA 890
QY 898 gaaattgttttgatttgggtggtctctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 957
Db 891 GAAATTTGTTTGAATTTGGGTGGTCTCTGTGGTGTCTGTGGCTGTCCCTGAAGCCATAGGTTT 950
QY 958 gatataattacgatggactacaagaagtaattctgcgaatctgtctgtgtgtgtgtgtgtgtgt 1017
Db 951 GATATAATTACGATGGACTACAAAGGAAGTTATCTCGGAATCTGCTTGTCTTCATCCCGTT 1010
QY 1018 cagaagacagctttcatgcagttttacaagacagcaaaaagattggtggtgttcaagtttc 1077
Db 1011 CAGAAGACAGCTTTTCATGTCAGTTTTTACAAGACAGCAAAAAGATTGGTGGCTATTTCAGTTTC 1070
QY 1078 tatttctgttgccttgccatcactgcatctgttttttttatcacactaatgacctgtgaaatg 1137

```

|||||
Db 1071 TATTTCTGCTGCCATTTGCCATCACTGATGATTTTATATACACTAATGACCTGTGAATG 1130
Qy 1138 ttgagaagaaaagtggaatgcagattgcttttaaatgatcaactaaagcagagacgggaa 1197
Db 1131 TTGAGAAAGAAAAGTGGCATGCAGATTGCTTTAAATGATCACTAAAGCAGAGACGGGAA 1190
Qy 1198 gtggccaaaacgcgtcttttgccgtgctcttgcctctgtgcttgcctctgtgcttcccttcac 1257
Db 1191 GTGGCCAAAACCGCTTTTGGCTGTCTTGTCTTGTGGCTGTGCTGCTGCCCTTCCCTTCAC 1250
Qy 1258 ctacagagattcgaactcactcttataatcagaatgatacccaatagatgtaactt 1317
Db 1251 CTCACAGGATTTCTGAAGCTCACTCTTATAATCAGATGATCCCAATAGATGTAACAT 1310
Qy 1318 ttgagctttctgttgattgactatatttggtatcaacatggtcttcaatcctgc 1377
Db 1311 TTGAGCTTCTGTTGGTATGACATATATGCTATCAACATGCTTCACTGAATTCCTGC 1370
Qy 1378 attaacccaattgctctgatttggtgagcaaaagattcaaaaactgctttaagtcacgc 1437
Db 1371 ATTAACCCAAATTCCTCTGTTATTTGGTGAGCAAAAGATTCAAAAACCTGCTTTAAGTCATGC 1430
Qy 1438 ttactgctgctgcagtcatttgaaagaaaacagtccttgagagaaaagcagtcgac 1497
Db 1431 TTATGCTGCTGTGCGCAGTCATTTGAAGAAAACAGTCCTTGGAGAAAAGCAGTCGTGC 1490
Qy 1498 ttaagttcaaaagtaatgatacaggaatgacacgttcctgctcagtaataatacagc 1557
Db 1491 TTAAAGTTCAAGCTAATGATCAGGATATGACAACTTCCGTTCAGATTAATAATACAGC 1550
Qy 1558 tcattctggaagaagaacttcaactgatttcattcttcttatttattgagccgaagtcat 1617
Db 1551 TCATCTTGAAGAAGAACTATTCACTGTATTTTCATTTCTTATATTGGACCGAAGTCAT 1610
Qy 1618 taaacaaataaagcatttgcaaaacaaacaaacaaactatgtatttgacacacac 1677
Db 1611 TAAACAAATGAACATTTGCCAAACAAACAAACAAACAACTATGATTTTGACAGCACAC 1670
Qy 1678 tattaataatlaagtgaattatttttaacactcacagctacatatgacattttatgagc 1737
Db 1671 TATTAATAATTAAGTGAATTAATTTTAACACTCACAGTACATATGACATTTTATGAGC 1730
Qy 1738 tgtttacggcatggaagaaaatcagtgggaaattaagaagcctcgtcgaagacatt 1797
Db 1731 TGTTTACGGCATGGAAGAAAATCAGTGGGAATTAAGAAAGCCTCGCTCGTGAAGACAT 1790
Qy 1798 aatttttacagttagcactcaacatacgtcttcaacactccaggatattcacacaac 1857
Db 1791 AATTTTTCAGTTAGCACTTCAACATAGCTCTTAAACAACTTCCAGGATATTACACAAC 1850
Qy 1858 acttaggcttaaaatgagctca 1880
Db 1851 ACTTAGCTTAAATAGACTCA 1873

RESULT 10
S44866 S44866 1872 bp mRNA linear PRI 07-MAY-1993
LOCUS ETB endothelin receptor [human, mRNA, 1872 nt].
DEFINITION S44866
ACCESSION S44866
VERSION S44866.1 GI:233233
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1872)
AUTHORS Sakamoto,A., Yanagisawa,M., Sakurai,T., Takuwa,Y., Yanagisawa,H.
and Masaki,T.
TITLE Cloning and functional expression of human cDNA for the ETB
endothelin receptor
JOURNAL Biochem. Biophys. Res. Commun. 178 (2), 656-663 (1991)

91315496
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 44866] from the original journal article.
This sequence comes from Fig. 1.
Location/Qualifiers
source
1. 1872
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. 1872
/gene="ETB endothelin receptor"
CDS
231..1559
/gene="ETB endothelin receptor"
/note="This sequence comes from Fig. 1"
/codon_start=1
/product="ETB endothelin receptor"
/protein_id="AAB19411.1"
/db_xref="GI:233234"
/translation="MOPPSLCGALVALVACGLSRINGFERGPPDPDRATPLLOTAE
INTVPSCLFVIGIIGNSTLLRIYKKNKMRGNPNILIASLALDGLLHVIDIPNVI
KLAEDWPFGAEMCKLVPIQKASVIGITVLSICALSIDRYAVASWSRIKGVPKWT
AVEIVLWVSWLAVPEAIGFDIITMDYKGSYLKICLLHPVAKTAHQFKTKADWW
LFSVFCPLAITAFFYTLMTCEMLRKSQMIALNDHLKORREKAVKTVFCLVLFAL
CWLPLHLSRIKLTLNNDPNRCLELSFLVLVDYIGINMASLASCINPIALYLVSKR
EKNFCKSLCCWCQSFEEKQSLKFKANDHGYDNFRSSNKYSSS"
BASE COUNT 489 a 434 c 438 g 511 t
ORIGIN

Query Match 39.3%; Score 1690.8; DB 9; Length 1872;
Best Local Similarity 99.6%; Pred. No. 3.9e-292;
Matches 1695; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 178 tgaacttgctcctgaaactgcggagcggccaccggagcgcctctctgagcaggtagcagc 237
Db 171 TCTCTCTAGCTCTGAACTGCGAGGCGCCACCGACGCCCTCTGACACAGTAGCAGC 230
Qy 238 atgcgcgcctccaagtctgtgcggacgcgcctggtgtgcctggtcttctgcctgcgcgc 297
Db 231 ATGCAGCCGCCCTCAAGTCTGTGCGACCGCGCCCTGGTTGCGCTGGTTCTTGCCTCGGC 290
Qy 298 ctgtcgcgagatctggggagagagagagcgttccgcgcctgacagggccactccgcttttg 357
Db 291 CTGTCCGGATCTGGGGAGAGAGAGAGCGCTTCCGCCCTGACAGGCCCACTCGCCTTTTG 350
Qy 358 caaacgcagagataatgacgcaccaccactaaagacctatgcccagggttcccaacgc 417
Db 351 CAACCGCAGAGATTAATGACGCCACCCTTAAGACCTTATGCCCAAGGTTCCAAGGCC 410
Qy 418 agtctggcggtgctgtggcactgcggaggtgctctaaaggagagacagcagagatct 477
Db 411 AGCTGGCGCGGTGCTGTGGACCTGCGGAGGTCCCTAAAGAGAGACAGGACGCGAGATCT 470
Qy 478 ccgcacgcacacatctccccctcccgctgccaagaccctcagatcgaagagacttc 537
Db 471 CCGCACGCAACCATCTCCCTCCCGTCCCAAGGACCCCATCGAGATCAAGAGAGACTTTC 530
Qy 538 aaatacatcaacacaggtgtgctgcctgctgtgttctgctgggagatcatcgggaactcc 597
Db 531 AAATACATCAACACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
Qy 598 acactctgagaattatctacaagaacaaagtgcagtcgcaaacgggtcccaatatctctgac 657
Db 591 ACACCTCTGAGAATTATCTACAAGAACAAAGTCATCGAAGACGGTCCCAATATCTTGTATC 650
Qy 658 gccagctggtcctctgggagaacctgctgcacatcgctcattgacatccctcatcaatgctac 717
Db 651 GCCAGCTTGGCTCTGGAGACCTGTGTCACATCGTCACTTATGACATCCCTATCAATGCTAC 710
Qy 718 aagctgctggcagagactggtccatttgagctgagatgtgagctgtaagctggtgctcttcata 777
Db 711 AAGCTGCTGGCAGAGACTGGCCATTTGGAGCTGAGATGTGTAAGCTGGTGGCTTTCATA 770

1. 1872
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene
1. 1872
/gene="ETB endothelin receptor"
CDS
231..1559
/gene="ETB endothelin receptor"
/note="This sequence comes from Fig. 1"
/codon_start=1
/product="ETB endothelin receptor"
/protein_id="AAB19411.1"
/db_xref="GI:233234"
/translation="MOPPSLCGALVALVACGLSRINGFERGPPDPDRATPLLOTAE
INTVPSCLFVIGIIGNSTLLRIYKKNKMRGNPNILIASLALDGLLHVIDIPNVI
KLAEDWPFGAEMCKLVPIQKASVIGITVLSICALSIDRYAVASWSRIKGVPKWT
AVEIVLWVSWLAVPEAIGFDIITMDYKGSYLKICLLHPVAKTAHQFKTKADWW
LFSVFCPLAITAFFYTLMTCEMLRKSQMIALNDHLKORREKAVKTVFCLVLFAL
CWLPLHLSRIKLTLNNDPNRCLELSFLVLVDYIGINMASLASCINPIALYLVSKR
EKNFCKSLCCWCQSFEEKQSLKFKANDHGYDNFRSSNKYSSS"

```
Qy 778 cagaagcctcgtgggaatcaactgtgtgagctatgtgtgtctgtgtctgtgagtagtatat 837
Db 771 CAGAAGCCTCGTGGGAATCACTGTGTGAGTCTATGTGCTGTGCTGAGTATTGACAGATAT 830
Qy 838 cgagctgtctctctgagtagaattaaaggaattgggttccaaatggagcagcagta 897
Db 831 CGAGCTGTGCTCTGAGGAGTAGAATTAAAGGAATTGGGGTTCCAAATGGACAGCAGTA 890
Qy 898 gaaattgtttgattgggtggtctctgtgtgtctgtgtgtctgtgtgtctgtgtgtgttt 957
Db 891 GAAATGTTTGAATTGGGTGGTCTCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGT 950
Qy 958 gataataacagatggactacaaaggaagtatctgtcgaatctgtgtgtctgtgtgttt 1017
Db 951 GATAATAATACAGATGGAGTACAAAGGAAGTATCTGCGAATCTGTGCTTTCATCCCGTT 1010
Qy 1018 cagaagacagcttcatgcaagcttttacaagacagacaaagattggtgctgttcaatttc 1077
Db 1011 CAGAAGACAGCTTTCATGCAAGTTCATCAAGACAGCAAAAGATTGTTGGCTATFTCAGTTTC 1070
Qy 1078 tatttctgttgcattggccatcactgcattttttttatcacataatgacctgtgaaatg 1137
Db 1071 TATTCTGTCTGCCATTGGCCATCACTGCATTTTATATACATTAATGACCTGTGAATG 1130
Qy 1138 ttgagaagaaagtggcatgagatgctttaaagatcatcacctaaagcagagacgggaa 1197
Db 1131 TTGAGAAGAAAGTGGCATGCAGATTGCTTTAAATGATACCTTAAAGCAGAGACGGGAA 1190
Qy 1198 gtggccaaaacgcttttggctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1257
Db 1191 GTGGCCAAAACGCTTTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1250
Qy 1258 ctgagcaggattctgaagctcaactctttataatcagaatgatcccaatagatgtgaactt 1317
Db 1251 CTCAGCAGGATCTGAAGCTCACCTTTTATATCAGATGATCCCAATAGATGTGAATTT 1310
Qy 1318 ttgagcttctgtgtgattggacatataatggatcaacatggcttcaactgaattccctgc 1377
Db 1311 TTGAGCTTCTGTGGTATTTGGACTATATTGCTATCAACATGGCTTCACTGAATTCCTGC 1370
Qy 1378 attaaccaattgtctgtattgttgagaaagattcaaaactgctttaaagtcatgc 1437
Db 1371 ATTAACCAATTTGCTGTGATTTGGTGAGCAAAAGATTCAAAAATCTGCTTTAAGTCATGC 1430
Qy 1438 ttatgtctgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 1497
Db 1431 TTATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1490
Qy 1498 ttaagttcaagctaataatgatcaggtatgacaaacttccgttccagtaataataacagc 1557
Db 1491 TTAAGTTCAAGCTAATGATACGGGATATGACAACTTCCGTTCCAGTAAATAATACAGC 1550
Qy 1558 tcatttgaagaagaactatctgtatttcttcttcttcttcttcttcttcttcttcttct 1617
Db 1551 TCATCTTTGAAGAAGAAGTATTCATCTGTATTTCTATTTCTTTATATTGGACCGAAGTCAT 1610
Qy 1618 taaacaaaaatgaacatttgcaaaaaaacaacaaataatgtatttgcacagcacac 1677
Db 1611 TAAACAAAATGAACATTTTGCACAAAACAAACAAACAAATATGTATTTTGCACAGCACAC 1670
Qy 1678 tattaataattaaagttaattattttaaactcacagctacatatgacatttttatgagc 1737
Db 1671 TATTAAATAATTAAAGTAAATTTAATAACATCACAGCTACATATGACATTTTATGAGC 1730
Qy 1738 tgtttacggcatggaaagaaataacagtggaattaaagaagcctcgctgtgaagacactt 1797
Db 1731 TGTTTACGGCATGGAAAGAAATACAGTGGGAATTTAAGAAAGCCCTCGCTGTGAAGACACTT 1790
Qy 1798 aattttttacagtagcacttcaactagctcttaacaaacttccagatatttcaacaac 1857
Db 1791 AATTTTACAGTAGCATTCAACATAGCTCTTAAACAACTTCCAGGATATTTCACAAAC 1850
Qy 1858 acttaggcttaaaaatgagctc 1879
```

```
Db 1851 ACTTAGGCTTAAAAATGAGCTC 1872
```

RESULT 11

```
HUMEDNRB
LOCUS Homo sapiens endothelin receptor type B (EDNRB) mRNA linear PRI 07-NOV-1994
DEFINITION Homo sapiens endothelin receptor type B (EDNRB) mRNA, complete cds.
ACCESSION L06623
VERSION L06623.1 GI:181958
KEYWORDS endothelin receptor; endothelin receptor type B.
SOURCE Homo sapiens (tissue library: lambda ZAPII) lung cDNA to mRNA.
ORGANISM Homo sapiens
MAMMALIA: Eutheria; Primates; Catarrhini; Hominoidea; Homo.
EUKARYOTA: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
MAMMALIA: Eutheria; Primates; Catarrhini; Hominoidea; Homo.
1 (bases 1 to 1719)
REFERENCE
AUTHORS Elshourbady, N.A., Korman, D.R., Wu, H.L., Sylvester, D.R., Lee, J.A.,
Nuthalaganti, P., Bergsma, D.J., Kumar, C.S. and Namb, I.P.
TITLE Molecular characterization and regulation of the human endothelin
receptors
JOURNAL J. Biol. Chem. 268 (6), 3873-3879 (1993)
MEDLINE 93179382
FEATURES
source
1..1719
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="Unassigned"
/tissue_type="lung"
/tissue_lib="lambda ZAPII"
1..1719
/gene="EDNRB"
234..1562
/gene="EDNRB"
/codon_start=1
/product="endothelin receptor type B"
/protein_id="AAA52342.1"
/db_xref="GI:181959"
/db_xref="GDB:G00-129-075"
/translation="MOPPPSLCGRALVALVLACGLSRWGEERGFPPORPPILOTAE
IMTPTKTLWPKGNSLASLAPAEVFKGDRTAGSPRTISPPCQGFIEKETFY
INTVSCFLVGLIGNSTLLRIYKKNMRNPILIASLALGDLHLHVIDIPINVY
KLAEADPFGEAKLVPIQKASVITVLSICALSIDIRAVASWSRIGVPKWT
AVEVLIWVSWLAWPAIGFDIITMDYKGYLCILCLHPVOKTAFQFYKTDWV
LFSYFCLPLAITAFFYTLMTCEMLRKKSGMOIALNDHLKORREKAVKVFCLVLVFL
CWLPLHLSRLKLLYLNQDNRCLELLSFLVLVDYIGINMASLNCINPIALYLVSKR
FKNCFKSLCCWCOSFEBKQSLKFKANDHGYDNFRSNKSYSS"
BASE COUNT 444 a 400 c 412 g 463 t
ORIGIN
Query Match 34.8%; Score 1495.4; DB 9; Length 1719;
Best Local Similarity 99.4%; Pred. No. 3.1e-257;
Matches 1522; Conservative 0; Mismatches 6; Indels 3; Gaps 2;
Qy 178 tgaacttggtctgaaactcgagcgccaccgagcctcttgagcaggtagcagc 237
Db 177 TGTCTCTAGGCTCTGAACCTGG--GCGCCACCGGACG-CTTCTGGACAGGTAGCAGC 233
Qy 238 atgcagcgcctccaaagtctgtgcgagcgccctgggtgctgctgttcttgcctgcgc 297
Db 234 ATGCAGCGCCTCCAAAGTCTGTGCGACGCGCCCTGGTGGCTGTTCTTGTGCTCGCGC 293
Qy 298 ctgtcgcggaatctgggagagagagagagagagagagagagagagagagagagagagag 357
Db 294 CTGTCGCGGATCTGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 353
Qy 358 caaacccagagataataagcaccacccaactaagacctatgcccagaggttcccaacgcc 417
Db 354 CAACCCGAGAGATTAATACCGCACCCACCTATAGACCTTATGCCCCAAGGGTTCACACGCC 413
Qy 418 agtctggtgcggtctgtgtgacacctgcgaggtgtgcctaaagagagagagagagagagat 477
Db 414 AGTCTGGCGCGTGTGTGGACCTGCGGAGGTGCTTAAGAGAGACAGGACGACGAGTCT 473
```


QY	312	gggagaggagagaggtctccgcctcgacagggccactccgcttttggcaaacccgacagat	371
Db	121	GGGAGAGGAGAGAGGCTCCGCCTGACAGGGCCACCTCCGCTTTGGCAACCCGACAGAT	180
QY	372	aatgaagccaccactaagacacttatggcccaagggtccaaagccaggtccgccacgacacat	431
Db	181	AATGAGCCACCACCTAAGACTTATGSCCCCAAGGGTTCCAAGCCAGTCTGGCGCGGTC	240
QY	432	gttgccacctgcgaggtgctctaaaggagacagacgagcaggtatccgccacgacacat	491
Db	241	GTTGGCACTGGGAGGTGCCCTAAAGGAGACGAGCCGGCAGGATCTCCGCGACCACT	300
QY	492	ctccccccccgtgccaagaccacatcgagatcaagagadacttcaaatatcatcaaac	551
Db	301	CTCCCTCCCGCTGCCAAGACCCATCGAGATCAAGGACATTTCAATATCATCAACAC	360
QY	552	ggttgtgtccctgtgtgtgttcgtgctgggatactcgggaactccacactcttgagaat	611
Db	361	GGTGTGTCTGCTTGTGTTCGTGCTGGGGATCATCGGGAATCTCCACACTTCTGAGAA	420
QY	612	tatctacaagaacaagtgcattgcgaaacgtgccaatctctgatactgcgcaggttggtct	671
Db	421	TATCTACAGAACAAAGTGCATGCCAAACGGTCCCAATATCTTGATGCCAGCTTGGCT	480
QY	672	gggagacctgtgcacatcgctattgacatccctatcaatgctcacaagctctggcaga	731
Db	481	GGGAGACCTGCTGCACATCGTCATGTGACATCCCTATCAATGCTACAGCTCTGGCAGA	540
QY	732	ggactggccatttgagctgagatggtgtaagctggtgctttccatcacgaaagccctcgt	791
Db	541	GGACTGGCCATTTGGAGCTGAGATGTAAAGCTGGTGCTTTCATACAGAAAGCCTCGT	600
QY	792	gggaatcaactgctgagctatgactctgactctgactatgacagatatcgagcttggtctc	851
Db	601	GGGAATCACTGCTGAGTCTATGTGCTGTAAGATATGACAGATATCCAGCTGTGTCTTC	660
QY	852	ttggagtgaataataaaggaattggggtccaaatggacagcagtagaataattgtttgat	911
Db	661	TTGGAGTAGAATTAAGGAATTTGGGTTCCAAATGACAGCAGTAGAATTTGTTGAT	720
QY	912	ttgggtggtctctgtggtctgctgctccctgaagccatagggttttggatataataacgat	971
Db	721	TTGGGTGGTCTGTGGTCTGCTGCTCCCTGAAGCCATAGGTTTGTGATATAATTACGAT	780
QY	972	ggactcaaaaggaaattatctcgaaatctgctgttcttcacccgttcagaagaacagcttt	1031
Db	781	GGACTACAAGGAAGTTATCTGCGAATCTGCTTGCTTCTATCCGTTTCAGAAAGACAGT	840
QY	1032	catgcagttttacaagacagaaaagattgggtggtgttccagtttctattctctgtgcc	1091
Db	841	CATGCAGTTTTACAAGACAGAAAAGATTGGTGGCTGTTCAGTTCTATTTCTGCTTGCC	900
QY	1092	atgggccatacactgaattttttttatacctaataagacctgtgaaatggtgagaagaagaag	1151
Db	901	ATTGGCCATCACATGCAATTTTTTATACATAATACCTGTGAAATGTTTGAGAAAGAAAG	960
QY	1152	tggcatgcagattgctttaaattgatcaactaaagcagacggaagtggccaaaacagt	1211
Db	961	TGSGATGCAGATTGCTTTAAATGATCACCTAAGCAGAGCGGAAGTGGCCAAAACCGT	1020
QY	1212	cttttgctgggtctgtgttttgccctctgctgggtctccctctcacctcagcaggaattct	1271
Db	1021	CTTTTGCTGGTCTGTGCTTTTGCTGCTGCTGCTTCCCTTCACTTCACCTCAGCAGATTCT	1080
QY	1272	gaagctcaactctttataatcagaatgatcccaatagatgtaacttttgagctttcttgtt	1331
Db	1081	GAAGCTCACTCTTATTAATCAGAAATGATCCCAATAGATGTGAACCTTTTGAGCTTTCTG	1140
QY	1332	ggattggactatattggtatcaacatggcttcactgaattctctgataaacccaattgc	1391
Db	1141	GGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATTCCTGCAATTAACCAATTGC	1200

Qy	1392	tctgtatttggtgagcaaaagattcaaaaaactgctttaagtcattgcttattgctgctggtg	1451
Db	1201	TTGTGATTTTGGTGAGCAAAAGATTCAAAACTGCTTTAAGTCATGCTTTAAGTCGTGGTG	1260
Qy	1452	ccagtcatttgaagaaaaaacagtccttgagagaaaagcagtcgctctaaagttcaaacg	1511
Db	1261	CCAGTCATTTGAAGAAAAACAGTCCTTGGAGGAAAAAGCAGTCGCTTTAAGTTCAAAGC	1320
Qy	1512	taatgatacagatgatgacaaactccggttccagtgataataatcacagctcatcttgaagaa	1571
Db	1321	TAAATGATCACGGATATGACAACTCCGTCCAGTAATAAATACAGCTCATCTTGAAGAA	1380
Qy	1572	gaactattcacgtgtattcattcttcattattgaccgaagtcattaaaaacaaatgaa	1631
Db	1381	GAACATATTCACGTGTATTTCAATTTTCTTTATATTGGACCGAAGTCATTAAAAACAAATGAA	1440
Qy	1632	acatttgcaaaacaaaaaactatg	1661
Db	1441	ACATTTGCCAAACAAAAACAAAAAACTATG	1470

RESULT 13

BC014472	BC014472	1603 bp	mRNA	linear	PRI 19-SEP-2001
LOCUS	Homo sapiens,				
DEFINITION	Similar to endothelin receptor type B, clone				
	MG:23118 IMAGE:486883,		mRNA, complete cds.		
ACCESSION	BC014472				
VERSION	BC014472.1		GI:15680234		

SOURCE: ORGANISM

ORGANISM

homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

1 (bases 1 to 1603)

AUTHORS

Submitted (17-SEP-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC/DCTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 34 Row: 1 Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4503466.

FEATURES

```

location/Qualifiers
1. .1603
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:23118 IMAGE:486863"
/tissue_type="Skin, melanotic melanoma, high MDR."
/clone_lib="NIH_MGC_49"
/lab_host="DH10B-R"
/note="vector: pOTB7"
236. .1564
CDS

```

cps

/codon_start=1
/product="similar to endothelin receptor type B"
/protein_id="AAH14472.1"
/db_xref="GI:15680235"
/translation="MQPPSLCGRALVALVACGLSRVGEERFPDPDRATPILQTAEL
INTPTTLKPKGNSALSLAEVPEKGRDTAGSPRTISPPPCQGPISIKETFKY
INTVYVLCVFLVGLIGNSTLRIYKKNMRNGPNILIASLALGDLHLVLDIPINVI
KLLAEDPFGAEMKLVFFIQKASVIGTILSCALSIDRYRAVASMRKIGVPKWT
AVETLVWVSVLVAPEAIGDITMDYKGSYLRICLLHPVKTAQFQKTKADWM
LFSVFLPLAITAFYFTLMTCEMLRKSGMOIALNDHLKORREKVAFTVCLVLFAL
CWLPLHSRLIKLTLNNDPNRCELFLVLDYIGINMASLNSCINPIALYLVSKR
FKNFKSLCCWCWCSFEKQSLBKFKANDHYDNFRSSNKYS"
BASE COUNT 407 a 379 c 399 g 418 t
ORIGIN

Query Match 32.3%; Score 1389; DB 9; Length 1603;
Best Local Similarity 99.6%; Pred. No. 3.1e-238; Indels 0; Gaps 0;
Matches 1392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 178 tgaacttgctctgaactcgagcgagcgccaccgagccctcttgagcaggtagcagc 237
Db 176 TGCTCTAGGCTCTGAAGCTCGGAGGCGCCACCGACGCCCTCTCGACGAGTAGCAGC 235
Qy 238 atgcagcgcctccaagtctgtgcggaagcgccctggctgctgctgctgctgctgctg 297
Db 236 ATGCAGCGCCCTCCAAAGCTGTGTGCGGACGCGCCCTGGTTCGCTGCTGCGCG 295
Qy 298 ctgtcgcgagatcgaggag 357
Db 296 CTGTCTGCGGATCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 355
Qy 358 caaacgcgagagataatgagcgcacccactaagacacttatggcccaaggttcccaagcc 417
Db 356 CAACCGCAGAGATAATGAGCGCCACCCACTAAGACCTTATGCGCCCAAGGTTCCCAAGGCC 415
Qy 418 agctcggcgctgtgtggaacctgcggaggtgctctaaaggagagagagagagagagatc 477
Db 416 AGCTGTGGCGCGTGTGTGCGACCTGCGGAGGTGCGCTAAAGGAGAGAGAGAGAGATC 475
Qy 478 ccgcacgcacacatccctccctcccggtgcgaagaccatcgagatcagagagacattc 537
Db 476 CCGCCACGACCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTTC 535
Qy 538 aaatacatcaacacaggtgtgtgctgctgctgctgctgctgctgctgctgctgctgctg 597
Db 536 AAATACATCAACAGGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 595
Qy 598 acactctgagagattatctacagaacaaagtcgcatgcgaaacggtcccaatactctgac 657
Db 596 ACACCTCTGAGAAATATCTACAGAACAAGTGCATCGCAACGGTCCCAATATCTTTGATC 655
Qy 658 gccactgtgctgagagacgtgcacatcgatcattgacatccctcatcattgctctac 717
Db 656 GCACGCTGGCTCTGGAGACCTGCTGCACATCGTCATTCGATCCCTATCAATGCTAC 715
Qy 718 agctgctgagagagactggccatttgagctgagatgtgtaagctgctgctgctgctgctg 777
Db 716 AAGCTGCTGGCAGAGACTGGCCATTTGGAGCTGAGATGTGTAAAGCTGCTGCTTTCATA 775
Qy 778 cagaagacccctcgtgggaatcactgtgctgagctctatgctgctgctgctgctgctgctg 837
Db 776 CAGAAAGCCCTCGGGAATCACTGTGCTGAGTCTATGCTGCTGAGTATTGACAGATAT 835
Qy 838 cgagctgtgcttcttgagtagaattaaaggaattgggtttcccaaatggagcagcagta 897
Db 836 CGAGCTGTGCTCTTGGAGTAGAATTAAGGAATTTGGGGTTCCTCAAAATGGACAGCAGTA 895
Qy 898 gaaattgtttgatttgggtggtctctgtggttctgctgctgctgctgctgctgctgctgct 957
Db 896 GAAATTGTTTGTATTTGGGTGGTCTCTGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT 955
Qy 958 gatataattacgatggacataaaggaagttatctgcggaatctgctgttcttccctccggt 1017

Db 956 GATATAATTACGATGAGTACAAAGGAAGTTATCTGCGAATCTGCTTCTCATCCGTT 1015
Qy 1018 cagaagacagcttctcatgagttttacaagacagcaaaagattggtgctgctgctgctgct 1077
Db 1016 CAGAAGACAGCTTTCATGAGCTTTTACAAGACAGCAAAAGATTGGTGGCTGTTTCACTTTC 1075
Qy 1078 tttcttctgcttgccttggccatcactgcatctttttttatcacataatgagacctgtaagt 1137
Db 1076 TATTCTGCTTCCCATTTGCCATCACTGCTATTTTATATACACTAATGACCTGTGAAATG 1135
Qy 1138 ttgagaagaaaagtggcagatgagattgctttaaattgatcacctcaaaagcagagagggaa 1197
Db 1136 TTGAGAAAAGAAAGTGGCATGAGATTGCTTTAAATGATCAGCTAAAGCAGAGACGGAA 1195
Qy 1198 gtggccaaaacccgtcttctgctgctgctgctgctgctgctgctgctgctgctgctgct 1257
Db 1196 GTGGCCAAAACCGCTCTTTTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1255
Qy 1258 ctacagcagagattctgaagctcactctttataatcagaatgatcccaaatagatgtaactt 1317
Db 1256 CTACAGCAGGATCTGAAGCTCACTCTTATATCAGAAATGATCCCAATAGATGTGAATTT 1315
Qy 1318 ttgagcttctgttggttggactatattggtatcaaatgagcttcaactgtaattcctgct 1377
Db 1316 TTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1375
Qy 1378 attaaccaaatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1437
Db 1376 ATTACCCCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1435
Qy 1438 ttatgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgctgct 1497
Db 1436 TTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1495
Qy 1498 ttaagttcaagcgaatgatcacgagatgacaaacttccgttccagtaataataacagc 1557
Db 1496 TTTAAAGTTCAAAGCTAATGATCAGGATATGACAACTTCGTTCCAGTAAATAATACAGC 1555
Qy 1558 tcactcttgaagaagaa 1574
Db 1556 TCATCTTGAAGAAAAA 1572

RESULT 14
AX342673
LOCUS Sequence 28 from Patent WO0198351.
DEFINITION
ACCESSION AX342673
VERSION AX342673.1 GI:18152053
KEYWORDS
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (sites)

REFERENCE
AUTHORS Lal,P., Baughn,M.R., Hafalia,A.J., Nguyen,D.B., Gandhi,A.R.,
Kallick,D.A., Griffin,J.A., Yue,H., Khan,F.A., Patterson,C.,
Lu,D.A., Tribouley,C.M., Lu,Y., Walia,N.K., Graul,R., Yao,M.G.,
Yang,J., Ramkumar,J., Au-Young,J., Hernandez,R., Walsh,R.T. and
Borowsky,M.L.
JOURNAL Patent: WO 0198351-A 28 27-DEC-2001;
Incyte Genomics, Inc. (US)

FEATURES
source
1. 1632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/note="Incyte ID No: 6792419CB1"
BASE COUNT 400 a 394 c 417 g 421 t
ORIGIN

Query Match 32.3%; Score 1389; DB 6; Length 1632;

Best Local Similarity 100.0%; Pred. No. 3.1e-238;			
Matches 1389;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy 186	ggctctgaactgcggagcgcgcacccgacacccctcttgagcaggtacgacatgcagcc	245	
Db 236	GGCTCTGAACCTGGGACGGCCACCGGACGCCCTTCTGGAGCAGGTACGACATGACGCC	295	
Qy 246	gcctccaaagtctgtgcggacgcgcctggttgcgtggttcttgcctgcgctctgtcgcg	305	
Db 296	GCCTCAAAGTCTGTGCGACGCGCCTGTTGCGTGTCTTGCTTGGCGCTGTGCGG	355	
Qy 306	gattctgggagagagagagaggtctccgccttgacagggccaactccgcttttgcgcaaacgc	365	
Db 356	GATCTGGGGAGGAGAGAGAGGCTTCCGCCCTGACAGGGCCACTCCGCTTTTGCAGAACGC	415	
Qy 366	agagataatgacgcccaccacctaagacctatgcccgaaggtttccaacgcagctctggc	425	
Db 416	AGAGATATGACGCCACCCACTAGACCTTATGCCCAAGGGTTCGAACGCCATCTGTGGC	475	
Qy 426	gcggtctgttgcaacctgcggaggtgccttaaggagacagagcgacaggtactccgcgcaacg	485	
Db 476	CGGGTCGTTGGCACTGCGGAGGTGCCTTAAAGGAGACAGGACGCGAGATCTCCGCCACG	535	
Qy 486	caccatctccctcccccgctgcgaaggaacccatcgagataaagagacatttcaaatatcat	545	
Db 536	CACCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTTCAAATACAT	595	
Qy 546	caacacggttgtctctgccttgttcgtgcgtgggacatcgcggaactccacactctct	605	
Db 596	CAACACGGTGTGTCCTGTCCTTGTTCGTGCTGGGGATCATCGGGAACATCCACATCTCT	655	
Qy 606	gagaattatctacaagaacaaagtgcacgcgaaacggtcccaatatcttgatcgccagctt	665	
Db 656	GAGAAATTATCTACAAAGAACAGTGCATGCGAAGCGTCCCAATATCTTGATCGCCACGTT	715	
Qy 666	ggctctgggagacgtgtgcacatcgcttgcattgacatccctatcaatgtctacaagctgt	725	
Db 716	GGCTCTGGGAGACGTGCTGCACATCGCATTCGATGACATCCCTATCAATGCTTACAAAGCTGCT	775	
Qy 726	ggcagaggaactggccatttgagagctgagatgtgaagctggtgccttttcatacagaagc	785	
Db 776	GGCAGAGACTGGCCATTTGGAGCTGAGATGTGTAAGCTGGTGGCTTTTCATACAGAAAGC	835	
Qy 786	ctcogtgggaactcactgtcgtgagctctatgtcctgagtattgacagatcatgcagctgt	845	
Db 836	CTCGTGGGAATCACTGTGCTGAGTCTATGTGCTCTGAGTATTGACAGATATCGAGCTGT	895	
Qy 846	tgctcttgagtagaattaaaaggaattggggttccaaaatgacagcagctagaaattgt	905	
Db 896	TGCTTCTTGAGTAGAATTAAGAGGAATTTGGGGTTCCTCAAAATGGACAGCAGTAGAAATGT	955	
Qy 906	tttgatttgggtggtctctgtgtctgcctgctgccttgcctgaagccatagggttttgataaat	965	
Db 956	TTTGATTGGGTGGTCTCTGTGGTTCTGGCTGTCCCTTGAAGCCATAGTTTTTGATATAAT	1015	
Qy 966	tacgattggactacaaaggaagtattctgcgaattctgttccttccctcccgcttcagaagac	1025	
Db 1016	TACGATGGACTACAAAGGAAGTTATCTGCGAATCTGCTTCTTCATCCCGTTCAGAAGAC	1075	
Qy 1026	agctttcatgcaattttacaagacagcaaaagattgggtgctgtcagtttctattctctg	1085	
Db 1076	AGCTTTTCATGACGTTTTTACAAGACAGCAAAAGATTGGTGGCTGTTTCAGTTTCTATTTCCTG	1135	
Qy 1086	cttggcaattggccatcactgcattttttatatacaactaatgacctgtgaaattgtgaaa	1145	
Db 1136	CTTGCCATGGCCATCAGTCGATTTTTTTATACACTAATGACCTGTGAATGTTTGAGAAA	1195	
Qy 1146	gaaaagtggcatgcagattgctttaaattgatcacctaaagcagagacgggaagtgcccaa	1205	
Db 1196	GAAGAGTGGCATGCAGATTGCTTTAATGATCACCTTAAGACGAGACGGGAAGTGGCCAA	1255	
Qy 1206	aaccgtttttgctggtctgtgttggcctctgtggtctcccttcaactcagcag	1265	

Db	1256	AACCGCTCTTTGGCCGTGCTTGTCTTTGGCCCTCTGCTGCTTCCCTTCACTCAGCAG	1315
Qy	1266	gattctgaagctcactcttataatcagaatgatcccaatagatgaaacttttgaagctt	1325
Db	1316	GATTCGGAAGCTCACTCTTTAATACGAATGATCCCAATAGATGGAATTTTGAAGCTT	1375
Qy	1326	tctgttggtattggactatttggtatcaaatggcttcactgaattcctgcattaaacc	1385
Db	1376	TCTGTTGGTATTGGACTATATTGGTATCAACATGCTTCACTGAATTCCTGCATTAACCC	1435
Qy	1386	aattgctctgatttggtagcaaaagatcaaaaactgctttaaagtcattatgctg	1445
Db	1436	AATTGCTCTGATTGTTGGTACAAAGATTCAAAACACTGCTTTAAGTCATGCTTATGCTG	1495
Qy	1446	ctgttgccagtcatttgaagaaaacagctccttgagagaaaagcagtcgctctaaagt	1505
Db	1496	CTGTGCCAGTCATTGGAAGAAAACAGTCTCTTGGAGGAAAAGCAGTCTGCTCTTAAGTT	1555
Qy	1506	caagctaatgatcacggatgatgaacattccgttcagtgataataatacagctcatcttg	1565
Db	1556	CAAGACTAATGATCAGCGATATGACAACCTCCGTTCCAGTATAATAATACAGCTCATCTTG	1615
Qy	1566	aaacaagaa	1574
Db	1616	AAAGAAGAA	1624
RESULT	15		
AF114165			
LOCUS	AF114165	1765 bp	linear
DEFINITION	Homo sapiens endothelin receptor B delta 3 mRNA		PRI 12-APR-1999
ACCESSION	AF114165		
VERSION	AF114165.1	GI:4580923	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 1765), Tsutsumi, M., Liang, G. and Jones, P.A.		
TITLE	Novel endothelin B receptor transcripts with the potential of generating a new receptor		
JOURNAL	Gene 228 (1-2), 43-49 (1999)		
MEDLINE	99173868		
REFERENCE	2 (bases 1 to 1765) Tsutsumi, M., Liang, G. and Jones, P.A.		
AUTHORS	Direct Submission		
TITLE	Submitted (17-DEC-1998) Urology, USC/Norris Cancer Center, 1411		
JOURNAL	Eastlake Ave., Los Angeles, CA 90033, USA		
FEATURES	Location/Qualifiers		
source	1. .1765 /organism="Homo sapiens" /db_xref="taxon:9606" 134. .1749 /codon_start=1 /product="endothelin receptor B delta 3" /protein_id="AAD24541.1" /db_xref="GI:4580924" /translation="MNKSTCLMAAETPSKRWRLHCLAFSQRFVRAGPACSSREACSSP RAGNAPRLRGWSPFVALHLVCOIREALKLRPPDASGAGSSMQPPSLGCRALV ALVACGLSRIGERGFPPDRATPLLOTAEIMTPPTKTLWPKGSNALSRLAPAEV PKGDRATGSPPTISPPPCQGPTEIKETFKYINTVVSCLVFLVGIIGNSTLLRIYKNI KCMRNGPNLTLASLAGDHLHVIDIPINVKLLAEDPFGAEMCKLVPIOKASGVI TVLSICALSIDRYAVASHRSRIGKIGVPKWTAVEIVLIWVSVLVAPPAIGFDIITM DKXGSYLRLCLHPVQKTFAMQFYKTAKDWMLFSFYFCLPLAITAFYFTMTCEMLRK KSGMIOALNDHLKSRKREAVTFCVLFCALCWLPLHLRLKLTLYLNDNDPNCRELL SFLVLDYTGINWASLNSINPIALVLSKRFKNGFKSCLCCWCQSFEEKQSLEEKQS CLKFKANDHGYDNFSSNKYS"		
BASE COUNT	425 a	434 c	442 t
ORIGIN			

2:

2:

2:

Search completed: October 1, 2002, 00:46:43
Job time: 27827 sec

34 sPArGAlaThrProLeuLeuGlnThrAlaGluIleMetThrProProThr 50
388 AAGACCTTATGGCCCAAGGGTTCCACGCCAGCTGTGGCGGTGGTGGC 437
51 LysThrLeuThrProLysGlySerAsnAlaSerLeuAlaArgSerLeuAl 67
438 ACCTCGGAGGTGCTTAAGGAGACAGACGCCAGGATCTCCGCCACGCA 487
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArgT 84
488 CCATCTCCCTCCCGCTGCCAAGGACCCATCGAGATCAAGGAGACTTTC 537
84 hrIleSerProProCysGlnGlyProIleGluIleLysGluThrPhe 100
538 AAATACATCAACACGGTGTGTCTGCCCTGTGTCTGGCTGGGAGATCAT 587
101 LysTyrIleAsnThrValValSerCysLeuValPheValLeuGlyIleI 117
588 CGGGAACCTCCACACTCTGAGAAATATCTACAAGAACAGTGCATGCCAA 637
117 eGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArg 134
638 ACGTCCCAATATCTGTATCCCGCAGCTGTGGCTCTGGGAGACTGTCGAC 687
134 snGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuHis 150
688 ATCGTCAATTGACATCCCTATCAATGTCTACAAAGCTGTGGCAGAGGACTG 737
151 IleValIleAspIleProIleAsnValTyrLysLeuLeuAlaGluAspTr 167
738 GCCATTTGGAGCTGAGATGTAAAGCTGGTGCCTTTTCATACAGAAAGCCT 787
167 pPropheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGFGGAACACACGTCTGAGTCTATGTCTGTCTGAGTATTCAGAGATAT 837
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
838 CGAGCTGTGCTCTTGGAGTAGAATTAAGGAATTTGGGTTTCCAAAGT 887
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
888 GACAGAGTAGAATTTGTTGATTTGGTGGTCTCTGTGGTCTGGCTG 937
217 pThrAlaValGluIleValLeuIleTrpValValSerValValLeuAlav 234
938 TCCTGAAGCATAGTGTGTGATATATACGATGGACTACAAAGGAGT 987
234 alProGluAlaIleGlyPheAspIleIleThrMetAspTyrLysGlySer 250
988 TATCTGCGAATCTGTGCTTCATCCCGTTTCAGAAAGACAGCTTTCATGCA 1037
251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetGl 267
1038 GTTTTACAGACAGCAAAAGATGGTGGCTGTTCAGTTTCTATTCTGCT 1087
267 nPheTyrLysThrAlaLysAspTrpLeuPheSerPheTyrPheCysL 284
1088 TGCATTGGCCATCACTGCATTTTTTTATACACTAATGACCTGTGAATG 1137
284 euProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGluMet 300
1138 TTGAGAAAGAAAGTGGCATGCGATGCTTTAAATGATCACCCTAAAGCA 1187
301 LeuArgLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGl 317
1188 GAGACGGGAGTGGCCAAACCGCTTTTGGCTGCTGTCTGTCTTGGCC 1237
317 nArgArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 334
1238 TCTGCTGGCTTCCCTTCCCTACCTCAGCAGGATCTGAAGCTCACTTTAT 1287
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyr 350

1288 AATCAGAATGATCCCAATAGATGCTGAACCTTTTCAGCTTCTGTGTGTTAT 1337
351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuValle 367
1338 GGACTATATTGGTATCAACATGCTTCACTGAATTCCTGCAATTAACCAA 1387
367 uAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnProI 384
1388 TTGCTCTGTATTGGTGAGCAAAAGATTCAAAAGCTGCTTTAAGTCATGC 1437
384 leAlaLeuTyrLeuValSerLysArgPheLysAsnSerCysPheLysSerCys 400
1438 TTATCTGCTGCTGCGCAGTCATTGTGAAGAAACACAGCTCCTTGGAGCAAA 1487
401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGluL 417
1488 CGAGTCGTGCTTAAAGTTCAAGCTTAATGATCAGCGATATGACAACCTCC 1537
417 sglnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnPheA 434
1538 GTTCAGTAATAATAACAGCTCATCT 1563
434 rgSerSerAsnLysTyrSerSerSer 442
seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep.US-08-117-361C-1
seq_documentation_block:
; Sequence 1, Application US/08117361C
; Patent No. 5773223
; GENERAL INFORMATION:
; APPLICANT: Venkatakrishna, Shyamala
; APPLICANT: Tekamp-Olson, Patricia
; TITLE OF INVENTION: Endothelin B1 (ETB1) Receptor Polypeptide
; TITLE OF INVENTION: Compositions, Methods, and Uses Thereof
; NUMBER OF SEQUENCES: 27
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Chiron Corporation
; STREET: 4560 Horton Street
; CITY: Emeryville
; STATE: California
; COUNTRY: USA
; ZIP: 946087
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 &
; SOFTWARE: Word 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/117,361C
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Chung, Ling-Pong
; REGISTRATION NUMBER: 36,482
; REFERENCE/DOCKET NUMBER: 0945.001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (510) 923-2704
; TELEFAX: (510) 655-3542
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 452 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY:
; LOCATION: 199..208
; OTHER INFORMATION: /note= "Decapeptide Insert"
US-08-117-361C-1

alignment_scores:

Quality: 2325.00 Length: 452
Ratio: 5.260 Gaps: 1
Percent Similarity: 97.788 Percent Identity: 97.788

alignment_block:

US-09-931-157-2 x US-08-117-361C-1 ..

Align seg 1/i to: US-08-117-361C-1 from: 1 to: 452

```
238 ATGCAGCGCGCTCAAGCTGTGCGGACGCGCCCTGGTGGCGTGTCT 287
1 MetGlnProProSerLeuCysGlyArgAlaLeuValLeu 17
288 TGCCTGCGCGCTGCGCGGATCTGGGAGAGAGAGAGGCTTCCCGCTG 337
17 uAlaCysGlyLeuSerArgIleTrpGlyGluGluArgGlyPheProA 34
338 ACAGGGCCTCCCTTTTGCACCCGACAGAGATAATGACGCCACCCACT 387
34 spArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProThr 50
388 AAGACCTATGGCCCAAGGFTCCACGCCAGTCTGCGCGGTGGTGGC 437
51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeu 67
438 ACCTGCGGAGTGCTAAAGAGAGAGAGCGGACGATCTCCGCCACGCA 487
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArg 84
488 CCATCTCCCTCCCGCTGCGAAGGACCCATCGAGATCAAGGACCTTTC 537
84 hrIleSerProProCysGlnGlyProIleGluIleLysGluThrPhe 100
538 AAATACATCAACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 587
101 LysThrIleAsnThrValSerCysLeuValPheValLeuGlyIleI 117
588 CGGGAATCCACACTCTGAGAATTATCTACAAGAACAACTGCATGCCAA 637
117 eGlyAsnSerThrLeuLeuArgIleIleTyLysAsnLysCysMetArg 134
638 ACGTCCCATATCTTATCGCCAGCTTGGCTCTGGGAGACCTCTGCAC 687
134 snGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuHis 150
688 ATCGCTATGACATCCCTATCAATGCTACAAGCTGCTGGCAGAGACTG 737
151 IleValIleAspIleProIleAsnValTyLysLeuLeuAlaGluAspTr 167
738 GCCATTTGGAGCTGAGATGTAAAGCTGGTGCCTTTCATACAGAAAGCCT 787
167 pPropheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGTGGGAATCACTGCTGCTGAGTCTATGCTGCTGAGTATTGAC..... 831
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspSerLeu 200
832 .....AGATATCGAGCTGTGCTTCTTCTGGAG 857
201 LysTyAsnSerIlePheIlePheArgTyArgAlaValAlaSerTrpSe 217
858 TAGAATTAAGGAATTTGGGTTCCAAATGGACAGCAGTAGAATTTGTT 907
217 rArgIleLysGlyIleGlyValProLysTrpThrAlaValGluIleVal 234
908 TGATTTGGTGGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
234 euIleTrpValValSerValValLeuAlaValProGluAlaIleGlyPhe 250
958 GATATAATACGAGTACTACAAGGAAGTTATCTGCGAATCTCTGCTGCT 1007
251 AspIleIleThrMetAspTyLysGlySerTyLysLeuArgIleCysLeuLe 267
```

```
1008 TCATCCCGTTCAGAAAGACAGCTTTCATGCGAGTTTACAAAGACACAAAG 1057
267 uHisProValGlnLysThrAlaPheMetGlnPheTyLysThrAlaLysA 284
1058 ATTGGTGGCTCTTCAGTTTCTATTCTGCTTGGCCATTTGGCCATCAGTCCA 1107
284 spTrpTrpLeuPheSerPheTyLysLeuProLeuAlaIleThrAla 300
1108 TTTTATATACACTAATGACCTGTGAATGTTGAGAAAGAAAGTGGCAT 1157
301 PhePheTyThrLeuMetThrCysGluMetLeuArgLysLysSerGlyMe 317
1158 GCAGATTGCTTTAAATGATCACCTAAAGACAGAGAGAGTGGCCAAA 1207
317 tGlnIleAlaLeuAsnAspHisLeuLysGlnArgGluValAlaLysT 334
1208 CCGTCTTTTGGCTGCTGCTTGTCTTTGCTGCTGCTGCTGCTGCTGCT 1257
334 hrValPheCysLeuValLeuValPheAlaLeuCysTrpLeuProLeuHis 350
1258 CTACAGCAGGATTCGAAGCTCACCTCTTATAATCAGAAATGATCCCAATAG 1307
351 LeuSerArgIleLeuLysLeuThrLeuTyAsnGlnAsnAspProAsnAr 367
1308 ATGTGAACCTTTGAGCTTCTGTTGGTATTGGACTATATTGGTATCAACA 1357
367 gCysGluLeuLeuSerPheLeuLeuValLeuAspTyIleGlyIleAsnM 384
1358 TGGCTTCACCTCAATTCCTGCAATTAACCAATTCCTGCTGCTGCTGCTG 1407
384 etAlaSerLeuAsnSerCysIleAsnProIleAlaLeuTyLysLeuValSer 400
1408 AAAAGATTCAAAACTGCTTAACTCATGCTTAACTGCTGCTGCTGCTGCTG 1457
401 LysArgPheLysAsnCysPheLysSerCysLeuCysCysTrpCysGlnSe 417
1458 ATTTGAAGAAACACAGCTCTTGGAGGAAAGACAGCTGCTGCTTAAAGTTCA 1507
417 rPheGluGluLysGlnSerLeuGluGluLysGlnSerCysLeuLysPheL 434
1508 AAGCTAATGATCAGGATATGACAACTCCGCTTCCAGTAAATAATACAGC 1557
434 ysAlaAsnAspHisGlyTyAspAsnPheArgSerSerAsnLysTySer 450
1558 TCATCT 1563
451 SerSer 452
```

seq_name: /cgn2_6/ptodata/2/iaa/5A_COMB.pep:US-08-118-270-41

seq_documentation_block:

; Sequence 41, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METHODS THEREOF
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/118,270
/ FILING DATE: 09-SEP-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/943,236
/ FILING DATE: 10-SEP-1992
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Townsend, Kevin G.
/ REGISTRATION NUMBER: 34,033
/ REFERENCE/DOCKET NUMBER: MURPHY-2A
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-628-5197
/ TELEFAX: 202-737-3528
/ TELEX: 248633

/ INFORMATION FOR SEQ ID NO: 41:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 350 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-118-270-41

alignment_scores:

Quality: 1318.50 Length: 274
Ratio: 4.938 Gaps: 3
Percent Similarity: 97.445 Percent Identity: 91.606

alignment_block:

US-09-931-157-2 x US-08-118-270-41

Align seg 1/1 to: US-08-118-270-41 from: 1 to: 350

745 GGAGCTGAGATGTAGCTGGTGGCTTCATACAGAAAGCCCGTGGG 794
|||||
79 GlycysArgLeucysArgLeuValProPheIleGlnLysAlaSerValGI 95
795 AATCAGCTGCTGAGCTATGTGCTGTGATGATGACAGATATCGAGCTG 844
|||||
95 yIleThrValLeuSerLeucysAlaLeuSerIleAspArgTyArgAlav 112
845 TTGCTTCTTGGAGTAAATTAAGGATTCGGGTTCACAAATGGACAGCA 894
|||||
112 alaIleSerIlePheIleGlyValProLysIleThrAla 128
895 GTAGAAATGTTTGTATTTGGTGGTCTGTGCTGCTGCTGCTGCTG 944
|||||
129 ValGluIleValLeuIleThrValValSerValValLeuAlaValProGI 145
945 AGCATAGGTTTGTATATATACGATGGACTACAAAGGAAGTTATCTGC 994
|||||
145 ualalleGlyPheAsp...ThrThrSerAspTyLysGlyLysProLeuA 161
995 GAATCTGCTGCTTCTCCCGTTCAGAAAGACAGCTTTCATGCAAGTTTAC 1044
|||||
161 rGValCysMetLeuAsnProPheGlnLysThrAlaPheMet...PheTy 176
1045 AAGACA...GCAAAAGATTGGTGGCTTCAGTTCTATCTGCTGCTGCC 1091
|||||
177 LysThrAlaAlaLysAspIlePhePheAlaPheTyPheCysLeuPr 193
1092 ATTGCCATCACTGCATTTTTTATACACTAATGACCTGTGAATCTGTA 1141
|||||
193 OleuAlaIleThrAlaIlePheTyThrLeuMetThrCysGluMetLeuA 210
1142 GAAAGAAAGTGGCATGAGATGCTTTAAATGATCACCTAAAGCAGAGA 1191
|||||
210 rGlyLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArg 226
1192 CGGGAAGTGGCAACCGCTCTTTGGCTGGTCTGCTGCTGCTGCTGCT 1241
|||||
227 ArgGluValAlaLysThrValPheCysLeuValPheAlaLeuCy 243

1242 CTGGCTTCCCTTCCCTCAGCAGGATTCTGAAGCTCAGCTCTTTATATC 1291
|||||
243 strLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyAspG 260
1292 AGAATGATCCCAATAGATGTGAACCTTTTGGAGCTTCTGTGGTATTGGAC 1341
|||||
260 InSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeuValLeuAsp 276
1342 TATATTGGTATCAACATGCTTCACCTGATGATTCCTGATTAACCAATTGC 1391
|||||
277 TyrIleGlyIleAsnMetAlaSerIleAsnSerCysIleAsnProIleAl 293
1392 TCTGTATTGGTGAGCAAAAGATTCAAAAGCTTTTAAAGCTTAACTGCTTAT 1441
|||||
293 aLeuTyIleValSerLysArgPheLysAsnCysPheLysSerCysLeuC 310
1442 GCTGCTGGTGGCAGTCATTGTAAGAAAAACAGTCCTTGGAGAAAAACGAG 1491
|||||
310 ysCysTrpCysGlnThrPheGluGluLysGlnSerLeuGluGluLysGln 326
1492 TCGTGTAAAGTTCAAAAGCTAATGATCAGGATATGACAACTTCGCTTC 1541
|||||
327 SerCysLeuLysPheLysAlaAsnAspHisGlyTyAspAsnPheArgSe 343
1542 CAGTAATAATACAGCTCATCT 1563
|||||
343 rSerAsnLysTyIleSerSerSer 350

seq_name: /cgn2_6/ptodata/2/iaa/PCTUS_COMB.pep:PCT-US93-08528-41

seq_documentation_block:

/ Sequence 41, Application PC/TUS9308528

GENERAL INFORMATION:

/ APPLICANT: New York University

/ TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN

/ NUMBER OF SEQUENCES: 348

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: BROWDY AND NEWMARK

/ STREET: 419 Seventh Street, N.W., Suite 300

/ CITY: Washington

/ STATE: D.C.

/ COUNTRY: USA

/ Zip: 20004

COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS

/ SOFTWARE: PatentIn Release #1.0, Version #1.25

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: PCT/US93/08528

/ FILING DATE: 09-SEP-1993

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 07/943,236

/ FILING DATE: 10-SEP-1992

/ ATTORNEY/AGENT INFORMATION:

/ NAME: Townsend, Kevin G.

/ REGISTRATION NUMBER: 34,033

/ REFERENCE/DOCKET NUMBER: MURPHY-2 PCT

/ TELECOMMUNICATION INFORMATION:

/ TELEPHONE: 202-628-5197

/ TELEFAX: 202-737-3528

/ TELEX: 248633

/ INFORMATION FOR SEQ ID NO: 41:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 350 amino acids

/ TYPE: amino acid

/ STRANDEDNESS: single

/ TOPOLOGY: linear

/ MOLECULE TYPE: peptide

/ PCT-US93-08528-41

alignment_scores:
Quality: 1318.50 Length: 274
Ratio: 4.938 Gaps: 3
Percent Similarity: 97.445 Percent Identity: 91.606

alignment_block:
US-09-931-157-2 x PCT-US93-08528-41 ..
Align seg 1/1 to: PCT-US93-08528-41 from: 1 to: 350

745 GGAGCTGAGATGTGTAAGCTGGTCCCTTCATACAGAAAGCCCTCGTGGG 794
|||||
79 GlyCysArgLeuCysArgLeuValProPheIleGlnLysAlaSerValGI 95
|||||
795 AATCACTGCTGCTGAGTCTATGCTCTGAGTATTGACAGATATGAGCTG 844
|||||
95 YileThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrArgAlav 112
|||||
845 TTGCTCTTGGAGTAGAATTAAGGAATGGGGTTCCAAATGACAGCA 894
|||||
112 aAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpThrAla 128
|||||
895 GTAGAAATGTTTGTGATTGGCTGCTCTGCTGCTGCTGCTGCCCTGA 944
|||||
129 ValGluIleValLeuIleTrpValValSerValValLeuAlaValProGI 145
|||||
945 AGCCATAGGTTTGTATATAATTACGATGGACTACAAAGAGTATTCTGC 994
|||||
145 uAlaIleGlyPheAsp...ThrThrSerAspTyrLysGlyLysProLeuA 161
|||||
995 GAATCTGCTGCTTCATCCGCTTCAGAGACAGCTTTCATGCAGTTTTAC 1044
|||||
161 rgValCysMetLeuAsnProPheGlnLysThrAlaPheMet...PheTyr 176
|||||
1045 AAGACA...GCAAGAAGATTGGTGGCTGTTCAGTTTCTATTCTGCTGCC 1091
|||||
177 LysThrAlaAlaLysAspTrpTrpLeuPheAlaPheTyrPheCysLeuPr 193
|||||
1092 ATTGGCCATCACTGCATTTTTTTATACACTAATGACCTGGAATGTGA 1141
|||||
193 oLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 210
|||||
1142 GAAAGAAAGTGGATGCAGTATGCTTTAAATGATCACCTTAAGCACAGA 1191
|||||
210 rgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArg 226
|||||
1192 CGGAAGTGGCAAAACCGCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTG 1241
|||||
227 ArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaLeuCy 243
|||||
1242 CTGCTTCCCTTCCCTCACCTCAGCAGGATTCCTGAAGCTCACTCTTTAATC 1291
|||||
243 strLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyrAspG 260
|||||
1292 AGAATGATCCCAATAGATGTCAACTTTTGACCTTTCTGCTGATTTGGAC 1341
|||||
260 lnSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeuValLeuAsp 276
|||||
1342 TATATTGGTATCAACATGGCTTCACTGAATTCCTGCATTAACCCAAATGC 1391
|||||
277 TyrIleGlyIleAsnMetAlaSerIleAsnSerCysIleAsnProIleAl 293
|||||
1392 TCTGTATTGGTGAGCAAAAGATTCAAAAAGTGGTTTAAAGTCATGCTTAT 1441
|||||
293 aLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCysLeuC 310
|||||
1442 GCTGCTGTCGATCATTTTCAAGAAACAGTCTCTGGAGAGAAAGCAG 1491
|||||
310 ysCysTrpCysGlnThrPheGluGluLysGlnSerLeuGluGluLysGln 326
|||||
1492 TCGTCTTAAAGTCAAGCTAATGATCACCAGATATGACAACTTCCTCGGTC 1541
|||||
327 SerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnPheArgse 343
|||||

1542 CAGTATAATAATACAGCTCATCT 1563
|||||
343 rSerAsnLysTyrSerSerSer 350
|||||

seq_name: /cgn2_6/ptodata/2/1aa/6B-COMB.pep:US-08-121-446-2

seq_documentation_block:
Sequence 2, Application US/08121446
Patent No. 6313276
GENERAL INFORMATION:
APPLICANT: IMURA, HIROO
APPLICANT: NAKAO, KAZUMA
TITLE OF INVENTION: A HUMAN ENDOTHELIN RECEPTOR
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSES: MORRISON & FORSTER,
STREET: 755 Page Mill Road
CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1018
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/121,446
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER: US 07/911,684
FILING DATE: 10-JUL-1992
ATTORNEY/AGENT INFORMATION:
NAME: CIOTTI, THOMAS E.
REGISTRATION NUMBER: 21,013
REFERENCE/DOCKET NUMBER: 29900-20324.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 813-5600
TELEFAX: (415) 494-0792
TELEX: 706141
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 427 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-121-446-2

alignment_scores:
Quality: 1240.00 Length: 359
Ratio: 3.987 Gaps: 5
Percent Similarity: 86.630 Percent Identity: 63.510

alignment_block:
US-09-931-157-2 x US-08-121-446-2 ..
Align seg 1/1 to: US-08-121-446-2 from: 1 to: 427

505 TGCCAGAGACCCATCGAGATCAAGAGACTTTCAATACATCACACGGT 554
|||||
69 CysProGlnGlnThrLysIleThrSerAlaPheLysTyrIleAsnThrVa 85
|||||
555 TGTGCTGCTGCTTGTGCTGCTGGGATCATCGGGAACCTCCACTTC 604
|||||
85 lIleSerCysThrIlePheIleValGlyMetValGlyAsnAlaThrLeuL 102
|||||
605 TGAGAAATATCTACAAGAACAAAGTGCATGCGAAACGGTCCCAATATCTTG 654
|||||
102 euArgIleIleIleTyrGlnAsnLysCysMetArgAsnGlyProAsnAlaLeu 118
|||||


```

1531 AACTTCGGTTCAGTAATAAATACACG 1557
      ::  |||||:::|||| |||
      417 ThrAspArgSerHisIysAspSer 425

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:US-08-118-270-40

seq_documentation_block:
; Sequence 40, Application US/08118270
; Patent No. 5508384
; GENERAL INFORMATION:
; APPLICANT: Murphy, Randall B.
; APPLICANT: Schuster, David I.
; TITLE OF INVENTION: POLYPEPTIDES OF G-COUPLED PROTEIN
; TITLE OF INVENTION: RECEPTORS, AND COMPOSITIONS AND METH
; NUMBER OF SEQUENCES: 348
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK
; STREET: 419 Seventh Street, N.W., Suite 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/118,270
; FILING DATE: 09-SEP-1993
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 07/943,236
; FILING DATE: 10-SEP-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Townsend, Kevin G.
; REGISTRATION NUMBER: 34,033
; REFERENCE/DOCKET NUMBER: MURPHY-2A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-628-5197
; TELEFAX: 202-737-3528
; TELEX: 248633
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 333 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-118-270-40

alignment_scores:
      Quality: 1166.00      Length: 342
      Ratio: 3.953          Gaps: 8
      Percent Similarity: 86.257      Percent Identity: 66.374

alignment_block:
US-09-931-157-2 x US-08-118-270-40 ..

Align seg 1/1 to: US-08-118-270-40 from: 1 to: 333

541 TACATCAACACGGTGTGTGCTCGCTTGTCTGCTGGGATCATCGG 640
|||||:::|||||:::|||||:::|||||:::|||||
1 TyrIleasnThrIleValSerCysLeuValPheValLeuGlyIleIleGI 17

591 GAACCTCCACACTTCTCAGAAATTATCTACAAGAACAAAGTGCATCGAAACG 640
|||||:::|||||:::|||||:::|||||:::|||||
17 YAsnSerThrLeuLeuArgIleIleIleTyrIysAsnLysCysMetArgAsnG 34

641 GTCCCAATATCTTGATGCGCCAGCTTCGGCTCTGGGAGACCTGTCGACATC 690
|||||:::|||||:::|||||:::|||||:::|||||
34 IyProAsnIleLeuIleAlaSerIleAlaLeuGlyAspLeuLeuHisIle 50

```



```
791 TGGGAATCACTGCTGAGTCTATGCTGCTGAGTATTGACAGATATCGA 840
|||||
81 aIGlyIleThrValLeuAsnLeuCysAlaLeuSerValAspArgTyrArg 97
|||||
841 GCTGTTGCTTCTGGAGTAGAATTAAGGAATGGGGTCCCAAAATGGAC 890
|||||
98 AlaValAlaSerTrpSerArgValGlnGlyIleGlyProLeuValTh 114
|||||
891 ACCAGTAGAATTTGTTTGGTGGTCTGCTGCTGGTTCGGCTGCTCC 940
|||||
114 rAlaIleGluIleValSerIleTrpIleLeuSerPheIleLeuAlaIleP 131
|||||
941 CTGAAGCCATAGCTTTGATATAATACGATGACTACAAAGGAAGTTAT 990
|||||
131 roGluAlaIleGlyPheTrpMetValProPheGluTyrLysGlyAlaGln 147
|||||
991 CTGCGAATCTGCTTCTTCATCCCTTACAGACACAGCTTTCATGAGTT 1040
|||||
148 HisArgThrCysMetLeuAsnAlaThrSerLysLeu.....Ph 160
|||||
1041 TTACAAGACAGCAAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
|||||
160 eTyrGlnAspValLysAspTrpIlePheGlyPheTyrPheLeu.... 175
|||||
1091 CATTGGCCATCATGCTATTTTATACACTAATGACCTGTGAATGTTG 1140
|||||
176 ..LeuValCysThrAlaIlePheTyrThrLeuMetThrCysGluMetLeu 191
|||||
1141 AGAAGAAAGCTGCG...ATCAGATTCTTAAATGATCACCCTAAACA 1187
|||||
192 AsnArgArgAsnGlySerLeuArgIleAlaLeuSerGluHisLeuLysG 208
|||||
1188 GAGACGGGAAGTGCCAAAACCGCTTTTGGCTGCTGCTGCTGCTGCTG 1237
|||||
208 nArgGluValAlaLysThrValPheCysLeuValIlePheAlaL 225
|||||
1238 TCTGCTGCTTCCCTTCACTCAGCAGGATTCGAAGCTCAGCTCTTTAT 1287
|||||
225 euCysTrpPheProLeuHisLeuSerArgIleLeuLysLysThrValTyr 241
|||||
1288 AATCAGATGATCCCATAGATGACTTTTGGCTGCTGCTGCTGCTGCTG 1337
|||||
242 AspGluMetAspThrAsnArgCysGluLeuLeuSerPheLeuLeuLeu 258
|||||
1338 GGACTATTGTTATCAACATGCTTCACTGAATTCCTGCTGCTGCTGCTG 1387
|||||
258 t...TyrIleGlyIleAsnThrAlaThrMet...SerCysIleAsnProI 273
|||||
1388 TTGCTCTGATTTGGTGAGCAAAAGATTCAAAACCTGCTTTAAGTCATGC 1437
|||||
273 leAlaLeuTyrPheValSerLysLysPheLysAsnCysPheGlnSerCys 289
|||||
1438 TTATGCTGCTGCTG...CAGTCATTTGAAGAAACACAGCTCTTGGAGGA 1484
|||||
290 LeuCysCysCysCysTyrGlnSerLysSerIleMetThrSerValProme 306
|||||
1485 AAAGCAGCTGCTGCTTAAAGTTCAAGCTCAATGAT...CAGGATATGACA 1531
|||||
306 tGlnGlyThrSerIleGlnTrpLysAsnHisGluGlnAsnAsnHisAsn 323
|||||
1532 ACTTCGGTCCAGTAAATAACAGC 1557
|||||
323 hrGluArgSerSerHisLysAspSer 331
|||||
```

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:us-07-937-609-17

seq_documentation_block:

; Sequence 17, Application US/07937609

; Patent No. 5319073

; GENERAL INFORMATION:

; APPLICANT: WANK, Stephen A.

; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF

```
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/937,609
; FILING DATE: 19920902
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/931,248
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 17:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: mouse gastrin-releasing peptide receptor
; US-07-937-609-17
```

alignment_scores:

Quality: 498.50 Length: 362
Ratio: 2.149 Gaps: 8
Percent Similarity: 64.088 Percent Identity: 33.702

alignment_block:

US-09-931-157-2 x US-07-937-609-17

Align seg 1/1 to: US-07-937-609-17 from: 1 to: 384

```
484 CGACCATCTCCCTCCCGTCCCAAGACCCTCGAGATCGAAGAGAC 533
|||||
25 GlnSerLeuSerProLysMetAspAsn.....TrpPheHisProG 39
|||||
534 TTTCAAATACATCAACACGGTTGTGTCTGCTGCTGCTGCTGCTGCTG 583
|||||
39 yPheIleTyrValIleProAlaValTyrGlyLeuIleValIleGlyL 56
|||||
584 TCATCGGGAACCTCCACACTTCTGAGATTATCTACAAGACAAGTGGATG 633
|||||
56 euIleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 72
|||||
634 CGAAACGGTCCCAATATCTGTATGATCGCCAGCTTGGCTGCGACACCTGCT 683
|||||
73 ArgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLe 89
|||||
684 GCATCTGCTCATTCGATCCCTATCAATGTCTACAAGCTCTGCGACAGG 733
|||||
```

```
89 uLeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspA 106
734 ACTGGCCATTGGAGCTGAGATCTGAAGTGGTGCCTTTTCATACAGAAA 783
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
106 rGrpLeuPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLeu 122
784 GCCTCCCTGGGAATCACTGCTGCTGAGTCTATGCTGCTGAGTATTGACAG 833
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAspAr 139
834 ATATCGAGCTGCTCTCTCTGGAGTAGAATTAAGGAATTGGGTTCCAA 883
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
139 gTyrLysAlaIleValArgProMetAspIleGlnAlaSerHisAlaLeuM 156
884 AATGCAGCAGCAGTAGAAATTTGTTTGGTGGTGGTCTCTGCTGGTCTG 933
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
156 eTylsIleCysLeuLysAlaAlaLeuIleTrpIleValSerMetLeuLeu 172
934 GCTCTCCCTCAAGCATAGCTTTTGATATAATTACGATGAGCTACAAAGG 983
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
173 AlaIleProGluAlaValPheSerAspLeuHisProPheHisValLysAs 189
984 AAGTTATCTCGAATC.....TGCTTGCTTCATCCCGTTCAAGAAGACAG 1027
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
189 pThrAsnGlnThrPheIleSerCysAlaProTyrProHisSerAsnGluL 206
1028 CTTTCATGCACTTTTCAACAGACAAAGATTGGTGGTCTTCAGTTTC 1077
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
206 euHisProLysIleHisSerMetAlaSer.....PheLeuVal 218
1078 TATTCTGCTGGCATGGCCATCACTGCATTTTATATACACTAATGAC 1127
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
219 PheTyrValIleProLeuAlaIleSerValTyrTyrPheIleAl 235
1128 CTGTGAATGTGTGAAGAAAGTGGCAGTCGAGATTGCTTTAAATGATC 1177
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
235 aArgAsnLeuIleGlnSerAlaTyrAsnLeuProValcIuGlyAsnIleH 252
1178 ACCTAAAGCAG.....AGACGGGAAGTGGCCAAACCGCTTTT 1215
    |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
252 isValLysLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeu 268
1216 TGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
269 ValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTy 285
1266 GATTCTGAAGCTCACTCTTTATATACAGATGATCCCAATAGATGTGAAC 1315
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
285 rLeuTyrArgSerTyrHisTyrSerGluValAspThrSer.....M 299
1316 TTTTGAGCTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1365
    :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
299 etLeuHisPhe.....ValThrSerIleCysAlaHisLeuLeuAlaPhe 313
1366 CTGAATTCCTGCATTAACCAATGCTGCTGTTGTTGTTGTTGTTGTTGTT 1415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
314 ThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSerPh 330
1416 CAAAACTGCTTTAAGTCA...TGCTTATGCTGCTGCTGCTGCTGCTGCT 1462
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
330 eArgLysGlnPheAsnThrGlnLeuLeuLeuCysGlnProGlyLeuMeta 347
1463 AAAAAACAGCTCTGGAGAAAGCAGTGTGCTGCTTAAAGTTCAAAAGCT 1512
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
347 snArgSerHisSerThrGlyArgSerThrThrCysMet..... 359
1513 AATGATCAGGATATGACAACTTCGTTCCAGTAAT 1548
    :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
360 .....ThrSerPheLysSerThrAsn 366

seq_name: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:US-08-029-170-17
seq_documentation_block:
```

```
56 euileGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 72
634 CGAAGAGTCCCAATATCTGTATGCGCCAGCTTGCTGCGGAGACCTGCT 683
73 ArgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLe 89
684 GCACATCGCTATTGACATCCCTATCAATGCTACAAAGCTGCTGGCAGAG 733
89 IleLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspA 106
734 ACTGCCATTTGGAGCTGAGATGCTAGAGTGGTGGCTTTCATACAGAAA 783
106 rgTrpLeuPheGlyargIleGlyCysLysLeuIleProPheIleGln 122
784 GCCTCCGGGGAATCACTGCTGAGTCTATGCTGCTGAGTATTGACAG 833
123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAsp 139
834 ATATCGAGCTGCTGCTTCTTGGAGTAGAATTAAGGAATTTGGGTCCA 883
139 gTyrLysAlaIleValArgProMetAspIleGlnAlaSerHisAlaLe 156
884 AATGCACAGCAGTAGAATTTGTTGATTTGGTGGTCTCTGTGTTCTG 933
156 etLysIleCysLeuLysAlaLeuIleIlePheValSerMetLeuLe 172
934 GCTGCTCCTGAAGCCATAGGTTTTCATATAATTTACGATGGACTACAA 983
173 AlaIleProGluAlaValPheSerAspLeuHisProPheHisValLys 189
984 AGTTATCTCGGAATC.....TCGTTGCTTCATCCGTTTCAGAGAGCAG 1027
189 pThrAsnGlnThrPheIleSerCysAlaProTyrProHisSerAsnGlu 206
1028 CTTTCATGACAGTTTTCACACACAGCAAGATTTGGTGGCTGTTTCAG 1077
206 euHisProLysIleHisSerMetAlaSer.....PheLeuVal 218
1078 TATTCTGCTGTCATTTGCGCATCTGCTGCTGCTGCTGCTGCTGCTG 1127
219 PheTyrValIleProLeuAlaIleIleSerValTyrTyrPheIleAl 235
1128 CTGTGAATGTTGAGAAGAAAGTGGCATGCGATGCTGCTGCTGCTGCT 1177
235 aArgAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIle 252
1178 ACTAAGACAG.....AGACGGAAGTGGCCAAACCGCTTTT 1215
252 isValLysGlnIleGluSerArgLysargLeuAlaLysThrValLeu 268
1216 TGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
269 ValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIle 285
1266 GATTCTGACGCTCACTTTTAAATCAGAAATGATCCCAATAGATGTAAC 1315
285 rLeuTyrArgSerTyrHisTyrSerGluValAspThrSer.....M 299
1316 TTTTGAGCTTTCTGTGGATTGGACTATATTGGTATCAACATGGCTTCA 1365
299 eLeuHisPhe.....ValThrSerIleCysAlaHisLeuLeuAlaPhe 313
1366 CTGAATTCCTGATTAACCAATGCTGCTGCTGCTGCTGCTGCTGCTG 1415
314 ThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSer 330
1416 CAAAACTGCTTTAAGTCA...TGCTATGCTGCTGCTGCTGCTGCTGCT 1462
330 eArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGlyLeuMet 347
1463 AAAAAACAGCTCTTGGAGGAAGACAGTCTGCTGCTGCTGCTGCTGCT 1512
347 snArgSerHisSerThrGlyArgSerThrThrCysMet..... 359
```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep: PCT-US92-02091-2

seq_documentation_block:
; Sequence 2, Application PC/TUS9202091
; GENERAL INFORMATION:
; APPLICANT: Battley Jr., James F.
; APPLICANT: Corlay, Martha H.
; APPLICANT: Feldman, Richard I.
; APPLICANT: Harkins, Richard N.
; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Edwin P. Ching
; STREET: 1501 Harbor Bay Parkway
; CITY: Alameda
; STATE: CA
; COUNTRY: USA
; ZIP: 94501
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/02091
; FILING DATE: 19920313
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/426,150
; FILING DATE: 24-OCT-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/533,659
; FILING DATE: 03-JUN-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Ching, Edwin P.
; REGISTRATION NUMBER: 34090
; REFERENCE/DOCKET NUMBER: A-0092C
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-266-7476
; TELEFAX: 415-266-7400
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 384 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US92-02091-2

alignment_scores:
Quality: 498.50 Length: 362
Ratio: 2.149 Gaps: 8
Percent Similarity: 64.088 Percent Identity: 33.702

alignment_block:
US-09-931-157-2 x PCT-US92-02091-2 ..
Align seg 1/1 to: PCT-US92-02091-2 from: 1 to: 384
484 CGCACCATCTCCCTCCCTCCCGTCCCAAGGACCCATCGAGATCAAGGAC 533
25 GlnSerLeuSerProProLysMetAspAsn.....TrpPheHisProG1 39
534 TTTCAAAATACATCAACACACGCTTGTGCTGCTGCTGCTGCTGCTGCTG 583
39 yPheIleTyrValIleProAlaValTyrGlyLeuIleIleValIleGlyL 56
584 TCATCGGGAACTCCACACTTCTGAGAATTATTCTACAAGAACAAAGTGCATG 633

```

56 eutleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 72
   ::::::::::::::::::::|::|::|
634 CGAAACGGTCCCAATATGTTGATCGCCAGCTTGGCTCTGGGAGACCTGCT 683
   ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
73 ArgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeuLe 89
   ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
684 GCACATCGTCATTGACATCCCTATCAATGTCTACAAAGCTCTGGCAGAGG 733
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
89 uLeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspA 106
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
734 ACTGGCCATTGGAGCTGAGATGTGTAAGTGGTGCCTTTCATACACAAA 783
   ||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
106 rgrtrLeuPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLeu 122
   ||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
784 GCCTCGCTGGGAATCACTGCTGAGTCTATGCTCTGAGTATTGACAG 833
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAspA 139
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
834 ATATCGAGCTGTTGCTTCTTGGAGTAGAATTAAGGAATGGGGTTCCAA 883
   ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
139 gTyrLysAlaIleValArgProMetAspIleGlnAlaSerHisAlaLeuM 156
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
884 AATGCACAGCAGTGAATAATGTTTGGATTGGTGGTCTCTGCTGTTCTG 933
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
156 etLysIleCysLeuLysAlaAlaLeuIleTrpIleValSerMetLeuLeu 172
   ||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
934 GCTGCTCCTGAAGCATTAGGTTTGTATATTAATACGATGGACTCAAGA 983
   ||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
173 AlaIleProGluAlaValPheSerAspLeuHisProPheHisValLysAs 189
   ||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
984 AAGTTATCTCGGAATC.....TCGTTGCTTCATCCCGTTCAGAAGACAG 1027
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
189 pThrAsnGlnThrPheIleSerCysAlaProTyrProHisSerAsnGluL 206
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1028 CTTTCATGCACGTTTACAGACACACAAAGATTGGTGGCTGTTCAAGTTC 1077
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
206 euHisProLysIleHisSerMetAlaSer.....PheLeuVal 218
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1078 TATTCTGCTGTCATGGCCATCCTGCTGCTGCTGCTGCTGCTGCTGCTG 1127
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
219 PheTyrValIleProLeuAlaIleIleSerValTyrTyrPheIleAl 235
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1128 CTGTGAATGTTGAGAAAGAAAGTGCATGCAGATTGCTTTAAATGATC 1177
   :::|::|::|::|::|::|::|::|::|::|::|::|::|::|
235 aArgAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIleH 252
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1178 ACCTAAACGAG.....AGACGGGAAGTGGCCCAACCCGCTCTTT 1215
   ||::|::|::|::|::|::|::|::|::|::|::|::|::|
252 IValLysLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeu 268
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1216 TGCCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1265
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
269 ValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTy 285
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1266 GATTCTGAAGCTCACTCTTTTATAATCAGATCATCCCAATAGATGTGAC 1315
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
285 rLeuTyrArgSerTyrHisTyrSerGluValAspThrSer.....M 299
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1316 TTTTCAGCTTTCTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 1365
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
299 etLeuHisPhe.....ValThrSerIleCysAlaHisLeuLeuAlaPhe 313
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1366 CTGAATTCCTGCATTAACCAATCTCTGTTATTTGGTGAGCAAAAGATT 1415
   ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
314 ThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSerPh 330
   ||||| |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1416 CAAAACATGCTTTAAATCA...TGCTTATGCTGCTGGTCCAGTCATTG 1462
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
330 eArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGlyLeuMetA 347
   |::|::|::|::|::|::|::|::|::|::|::|::|::|::|
1463 AAGAAAAACAGCTCTTGGAGGAAAAACAGCTCGTCTTAAGTTCAAAGCT 1512
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
347 snArgSerHisSerThrGlyArgSerThrThrCysMet..... 359
1513 AATGATCAGGATATGACAACTTCGTTCCAGTAAT 1548
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|
360 .....ThrSerPheLysSerThrAsn 366

seq_name: /cqn2_6/ptodata/2/1aa/PCTUS_COMB.pep.PCT-US92-02091-8

seq_documentation_block:
: Sequence 8, Application PC/TUS9202091
: GENERAL INFORMATION:
: APPLICANT: Battey Jr., James F.
: APPLICANT: Corjay, Martha H.
: APPLICANT: Feldman, Richard I.
: APPLICANT: Harkins, Richard N.
: TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
: NUMBER OF SEQUENCES: 8
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Edwin P. Ching
: STREET: 1501 Harbor Bay Parkway
: CITY: Alameda
: STATE: CA
: COUNTRY: USA
: ZIP: 94501
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: PCT/US92/02091
: FILING DATE: 19920313
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/426,150
: FILING DATE: 24-OCT-1989
: PRIOR APPLICATION DATA: US 07/533,659
: APPLICATION NUMBER: US 07/533,659
: FILING DATE: 05-JUN-1990
: ATTORNEY/AGENT INFORMATION:
: NAME: Ching, Edwin P.
: REGISTRATION NUMBER: 34090
: REFERENCE/DOCKET NUMBER: A-0092C
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 415-266-7476
: TELEFAX: 415-266-7400
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 390 amino acids
: TYPE: AMINO ACID
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: PCT-US92-02091-8

alignment_scores:
  Quality: 491.00      Length: 367
  Ratio: 2.173        Gaps: 8
  Percent Similarity: 61.580  Percent Identity: 32.153

alignment_block:
US-09-931-157-2 x PCT-US92-02091-8 ..
Align seg 1/1 to: PCT-US92-02091-8 from: 1 to: 390

424 CGCGCGGTGGTGGCACCTCGGAGGTGCTTAAAGGAGACAGACGCGCAGG 473
3 SerLysSerLeuSerAsnLeuSerValThrThrGlyAlaAsnGluSerG1 19
474 ATCTCCGCCA.....CGCACCATCTCCCTCCCGTGCCTGCCAAAG 511
19 ySerValProGluGlyTrpGluArgAspPheLeuProAlaSerAspGlyT 36
```

512 GACCATCGAGATCAAGGAGACTTTTCAAATACATCAACACGGTGTGTGTC 561
 |||:|
 36 hrThrThrGluLeu.....ValleArg 43
 562 TGC.....CTGTGTTCTGCTGGGATCATCGG 590
 |||
 44 CysValleProSerLeuTyrrLeuLeulleThrValGlyLeuLeuGl 60
 591 GAATCCACACTCTTGAGAATTTATCTCAAGAACAAGTGCATGCCAAGC 640
 |||
 60 yAsnMetLeuValLysIlePheIleThrAsnSerAlaMetArgSerV 77
 641 GTCCCAATATCTTGATCCCGAGCTGGCTCTGGAGACCTGCTGCACATC 690
 |||
 77 alProAsnIlePheIleSerAsnLeuAlaGlyAspLeuLeuLeu 93
 691 GTCAATGCATCCCTATCAATGTCTACAAGCTGTCGCAGAGGAGTGC 740
 |||
 94 LeuThrCysValProValAspAlaSerArgTyrrPhePheAspGluTrpMe 111
 741 ATTTGGAGCTGAGATGTGAAGCTGGTCTTCATACAGAAACCTCCG 790
 |||
 110 tPheGlyLysValGlyCysLysLeuIleProValIleGlnLeuThrSerV 127
 791 TGGGAATCACTGTCTCAGTCTATGTGCTCTGAGTATTGACAGATATCGA 840
 |||
 127 alGlyValSerValPheThrLeuThAlaLeuSerAlaAspArgTyrrArg 143
 841 GCTGTGTCTTCTGGAGTAGAATTAAGGAATTTGGGTTCCAAATGGAC 890
 |||
 144 AlaIleValAsnProMetAspMetGlnThrSerGlyAlaLeuLeuLeuArgTh 160
 891 AGCAGTAGAAATTTGTTGATTGGTGGTCTCTGTGTTCTCGCTGCTCC 940
 |||
 160 rCysValLysAlaMetGlyIleTrpValValSerValLeuLeuAlaValP 177
 941 CTGAGCCATAGTTTGTATTAATTTACGATG...GACTACAAAGGAAT 987
 |||
 177 roGluAlaValPheSerGluValAlaArgIleSerSerLeuAspAsnSer 193
 988 TATCTCGCAATCTGCTGCTCATCCGTTACAGACAGACTTTCATGCA 1037
 |||
 194 SerPheThAlaCysIleProTyrrProGlnThrAspGluLeuHisProLy 210
 1038 GTTTTACAAGACAAAGAATTTGGTGGCTCTCAGTTCTATTCTGCT 1087
 |||
 210 sIleHisSerValLeu.....IlePheLeuValTyrrPheLeuI 223
 1088 TGCCATTGGCCATCACTCCATTTTTTATACACTAATGACCTGTGAATG 1137
 |||
 223 leProLeuAlaIleSerIleTyrrTyrrHisIleAlaLysThrLeu 239
 1138 TTGAGAAGAAAGTGGCATCGAGATGCTTTTAATGATGACCTAAAGCA 1187
 |||
 240 IleLysSerAlaHisAsnLeuProGlyGluTyrrAsnGluHisThrLysLy 256
 1188 GAGACGGGAAGTGCCCAAAACCGCTTTTGGCTGGTCTCTCTGCTTGC 1236
 |||
 256 sGlnMetGluThrArgLysArgLeuAlaLysIleValLeuValPheValG 273
 1237CTCTGCTGGCTCCCTTCACCTCAGCAGGATTTGACCTT 1325
 |||
 273 lCysPheIlePheCysTrpPheProAsnHisIleLeuTyrrMetTyrrArg 289
 1276 CTCACCTTTTATATCAGAAATGATCCCAATAGATGTGAACCTTTGACCTT 1325
 |||
 290 SerPheAsnTyrrAsnGluIleAspProSerLeuLeuGlyHisMetIleValTh 306
 1326 TCTGTG.....GTATGGACTATTGCTATCAACATGGCTTCACTGA 1369
 |||
 306 rLeuValAlaArgValLeuSerPhe.....GlyA 316
 1370 ATTCTCTGCATTAACCAANTGCTCTGCTATTGTTGGTGAAGAAAAGATTCAA 1419

```

|||||.....
316 snSerCysValanProPheAlaLeuTyrLeuLeuSerGluSerPheArg 332
|||||.....
1420 AACTGCTTTAAGTCATGCTATCTGCTGCTGGTCCAGTCATTTGAAGAAAA 1469
::: |||||.....
333 ArgHisPheAsnSerGlnLeuCysGlyArgLysSerTyrGlnGluAr 349
1470 A 1470
:
349 g 349

seq_name: /cgn2_6/ptodata/2/1aa/5A_COMB.pap:US-07-937-609-18

seq_documentation_block:
; Sequence 18, Application US/07937609
; Patent No. 5319073
; GENERAL INFORMATION:
; APPLICANT: WANK, Stephen A.
; TITLE OF INVENTION: CLONING AND FUNCTIONAL EXPRESSION OF
; TITLE OF INVENTION: CHOLECYSTOKININ RECEPTOR-ENCODING DNA
; NUMBER OF SEQUENCES: 29
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/07/937,609
; FILING DATE: 19920902
; CLASSIFICATION: 436
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/831,248
; FILING DATE: 07-FEB-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/861,769
; FILING DATE: 01-APR-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/928,033
; FILING DATE: 11-AUG-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 40399/166 NIHD
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 18:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 390 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: rat neuromedin B receptor
-07-937-609-18

```

```

alignment_scores:
  Quality: 489.50      Length: 363
  Ratio: 2.205         Gaps: 8
  Percent Similarity: 61.157  Percent Identity: 31.680

alignment_block:
  US-09-931-157-2 x US-07-937-609-18  ..

Align seg 1/1 to: US-07-937-609-18 from: 1 to: 390

```


Quality: 489.50 Length: 363
Ratio: 2.205 Gaps: 8
Percent Similarity: 61.157 Percent Identity: 31.680
alignment_block:
US-09-931-157-2 x US-08-029-170-18 ..

Align seg 1/1 to: US-08-029-170-18 from: 1 to: 390

```
478 CGGCACGACGACCATCCCTCCCTCCCGTCCCAAGGACCCATCGAGATCAA 527
|||||
2 ProProArgSerLeu.....ProAsnLeuSerLeuProThrGluAlaSe 16
528 GGAGACT..... 534
|||||
16 rgluSerGluLeuGluProGluValTrpGluAsnAspPheLeuProAsps 33
535 .....TTCAATATACATCAACACGGTTGTCTCTGC..... 564
33 erAspGlyThrAlaGluLeuValIleArGcysValIleProSerLeu 49
565 .....CTTGTCGTCTGGGATCATCGGAACGGTCCCAATATCTTGATCG 658
|||||
50 TyrLeuIleIleSerValGlyLeuLeuGlyAsnIleMetLeuValLy 66
569 AATTATCTACAGAACAGTGCATCGGAACGGTCCCAATATCTTGATCG 658
|||||
66 silePheLeuThrAsnSerThrMetArgSerValProAsnIlePheIle 83
659 CCAGCTTGCTCGGGAGACCTGCTGCACATCGTCATTCATCCCTATC 708
|||||
83 erAsnLeuAlaIleGlyAspLeuLeuLeuLeuThrCysValProVal 99
709 AATGCTCTACAGCTGCTGGCAGAGGACTGCCATTCGAGCTGAGATGTG 758
|||||
100 AspAlaSerArgTyrPheAspGluTrpValPheGlyLysLeuGlyCy 116
759 TAAGCTGGTCCCTTCATACAGAACGCTCCGTTGGGAATCACTGTGCTGA 808
|||||
116 slysLeuIleProAlaIleGlnLeuThrSerValGlyValSerValPhe 133
809 GTCTATGTCTGCTGATGTGACAGATATCGACCTGCTGCTTCTTGGAGT 858
|||||
133 hrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsnProMet 149.
859 AGAATTAAAGGAATTGGGCTTCCAAATGGACAGCAGTAGAATTTT 908
|||||
150 AspMetGlnThrSerGlyValValLeuThrSerLeuLysAlaValGl 166
909 GATTTGGTGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 958
|||||
166 yileTrpValValSerValLeuLeuAlaValProGluAlaValPheSerG 183
959 ATATAATTACGATG...GACTCAAGAGGAGTATCTCGGAATCTGCTTG 1005
|||||
183 luValAlaArgIleGlySerSerAspAsnSerSerPheThrAlaCysIle 199
1006 CTTTCATCCCGTTCAGAACAGCTTTCATGTCAGTTTACAGACACAGCAA 1055
|||||
200 ProTyrProGlnThrAspGluLeuHisProLysIleHisSerValLeu... 215
1056 AGATTGGTGGTGTTCAGTTCTTATTTCTGCTGCTGCTGCTGCTGCTGCT 1105
|||||
216 .....IlePheLeuValTyrPheLeuIleProLeuValIleIle 229
1106 CATTTTATTATACATATGACCTGTGAATGTTGAGAAAGAAAGTGGC 1155
|||||
229 erIleTyrTyrTyrHisIleAlaLysThrLeuIleArgSerAlaHisasn 245
1156 ATGACGATGCTTTAAATGATCACCCTAAAGCAG.....AGACG 1193
|||||
246 LeuProGlyGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLy 262
```

```
1194 GGAAGTGGCCAAACCGTCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1243
|||||
262 sArgLeuAlaLysIleValLeuValPheValGlyCysPheValPheCyst 279
1244 GCCTTCCCTTCCACCTCAGCAGGATTCAGAGCTCAGCTCTTTTATAATCAG 1293
|||||
279 rpPheProAsnHisIleLeuTyrLeuTyrArgSerPheAsnTyrLysGlu 295
1294 AATGATCCCAATAGATGTGAACCTTTTGGCTGCTGCTGCTGCTGCTGCT 1337
|||||
296 IleAspProSerLeuGlyHisMetIleValThrLeuValAlaArgValle 312
1338 GCACATATATTGCTATCAACATGGCTTCACTGAATTCCTGCATTAAACCAA 1387
|||||
312 uSerPhe.....SerAsnSerCysValAsnProp 322
1388 TTGCTCTCTATTGCTGAGCAAAAGATTCAAAAGCTGCTTTTAAAGTCATGC 1437
|||||
322 heAlaLeuTyrLeuLeuSerGluSerPheArgLysHisPheAsnSerGln 338
1438 TTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1476
|||||
339 LeuCysCysGlyGlnLysSerTyrProGluArgSerThr 351
```

seq_name: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US92-02091-6

seq_documentation_block:

; Sequence 6. Application PC/TUS9202091

; GENERAL INFORMATION:

; APPLICANT: Battey Jr., James F.

; APPLICANT: Corjay, Martha H.

; APPLICANT: Feldman, Richard I.

; APPLICANT: Harkins, Richard N.

; TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES

; NUMBER OF SEQUENCES: 8

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Edwin P. Ching

; STREET: 1501 Harbor Bay Parkway

; CITY: Alameda

; STATE: CA

; COUNTRY: USA

; ZIP: 94501

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US92/02091

; FILING DATE: 19920313

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/426,150

; FILING DATE: 24-OCT-1989

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 07/533,659

; FILING DATE: 05-JUN-1990

; ATTORNEY/AGENT INFORMATION:

; NAME: Ching, Edwin P.

; REGISTRATION NUMBER: 34090

; REFERENCE/DOCKET NUMBER: A-0092C.

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 415-266-7476

; TELEFAX: 415-266-7400

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 390 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; PCT-US92-02091-6

alignment_scores:

Quality: 489.50 Length: 363
Ratio: 2.205 Gaps: 8
Percent Similarity: 61.157 Percent Identity: 31.680

alignment_block:

US-09-931-157-2 x PCT-US92-02091-6

Align seg 1/1 to: PCT-US92-02091-6 from: 1 to: 390

478 CGCCAGCAGCACCATCCCTCCCTCCGCGCCAGGAGCCATCGAGTCAA 527
|||||
2 ProProArgSerLeu.....ProAsnLeuSerLeuProThrGluAlaSe 16
528 GGAGACT..... 534
16 rgluSerGluLeuGluProGluValTrpGluAsnAspPheLeuProAsps 33
535TTCAATACATCAACACGGTTGCTCTGCTGC..... 564
33 erAspGlyThrThrAlaGluLeuValIleArgCysValIleProSerLeu 49
565CTGTGCTTCTGCTGGGATCATCGGGAACCTCCACACTTCTGAG 608
50 TyrLeuIleIleIleSerValGlyLeuLeuGlyAsnIleMetLeuVally 66
609 AATTATACAAAGAACAGTGCATCGGAACGGTCCCAATATCTTGATCG 658
66 silePheLeuThrAsnSerThrMetArgSerValProAsnIlePheIleS 83
659 CCAGCTTGGCTGGGACCTGCTGCACATCGTCACTTGCATCCCTATC 708
83 erAsnLeuAlaIleGlyAspLeuLeuLeuLeuLeuLeuLeuLeuLeu 99
709 AATGCTCTACAAGCTGCTGCAGACCTGCGCCATTTGGAGCTGAGATGTG 758
100 AspAlaSerArgTyrPhePheAspGluTrpValPheGlyLysLeuGlyCy 116
759 TAAGCTGTGCTCTTATACAGAAAGCTCCGCTGGGAATCACTGTGCTGA 808
116 slusLeuIleProAlaIleGlnLeuThrSerValGlyValSerValPheT 133
809 GTCTATGCTCTGAGTATTGACATATCGAGCTGTTGCTCTTGGAGT 858
133 hrLeuThrAlaLeuSerAlaAspArgTyrArgAlaIleValAsnProMet 149
859 AGAATTAAGGAATTTGGGTTCCAAAATGGACAGCAGTAGAAATGTTTT 908
150 aspMetGlnThrSerGlyValValLeuTrpThrSerLeuLysAlaValcl 166
909 GATTGGTGGTCTCTGTGGTCTCGCTCCCTGAAGCCATAGCTTTTG 958
166 yileTrpValValSerValLeuLeuAlaValProGluAlaValPheSerG 183
959 ATATAATTACAGT...GACTACAAGGAAGTTATCTCGAATCTGCTTG 1005
183 luValAlaArgIleGlySerSerAspAsnSerSerPheThrAlaCysIle 199
1006 CTTTCATCCGTTTCAGACAGCAGCTTTCATGAGTTTTCAGACAGCAA 1055
200 ProTyrProGlnThrAspGluLeuHisProLysIleHisSerValLeu... 215
1056 AGATTGGTGGTGTTCAGCTTTCTATTCTGCTTGCCTTGGCCATCACTG 1105
216IlePheLeuValTyrPheLeuIleProLeuValIleIleS 229
1106 CATTTTTTTATACATAATGACCTGTGAATGTTGAGAAAGAAAGTGGC 1155
229 erIleTyrTyrThrHisIleAlaLysThrLeuIleArgSerAlaHisAsn 245
1156 ATGCAGATTGCTTTAAATCATCACCTAAAGCAG.....AGACG 1193
246 LeuProGlyGluTyrAsnGluHisThrLysLysGlnMetGluThrArgLy 262

1194 GGAAGTGCACCAAAACCGTCTTTTGGCTGGTCTCTCTTTCCTCCCTCTCT 1243
262 sArgLeuAlaLysIleValLeuValPheValGlyCysPheValPheCyst 279
1244 GGCCTCCCTTCACCTCAGCAGGATTCGAGCTCCTCTTTTATATATCAG 1293
279 rpPheProAsnHisIleLeuTyrLeuTyrArgSerPheAsnTyrLysGlu 295
1294 AATGATCCCAATAGATGTGAACCTTTTGGCTTCTGTTG.....GTATT 1337
296 IleAspProSerLeuGlyHisMetIleValThrLeuValAlaArgValle 312
1338 GACATATATTTGATCAACATGGCTTCACTGAATTCCTGCATTAACCAA 1387
312 uSerPhe.....SerAsnSerCysValAsnProp 322
1388 TTGCTCTGATTTGGTGAGCAAAAGATTCAAAAAGCTTTTAAGTCATGC 1437
322 heAlaLeuTyrLeuLeuSerGluSerPheArgLysHisPheAsnSerGln 338
1438 TTATGCTGCTGGTGCAGTCACTTTGAAGAAAAACAGTCC 1476
339 LeuCysCysGlyGlnLysSerTyrProGluArgSerThr 351

seq_name: /cgn2_5/ptodata/2/1aa/PCTUS_COMB.pep:PCT-US92-02091-4

seq_documentation_block:

Sequence 4, Application PC/TUS9202091
GENERAL INFORMATION:
APPLICANT: Battley Jr., James F.
APPLICANT: Corjay, Martha H.
APPLICANT: Feldman, Richard I.
APPLICANT: Harkins, Richard N.
TITLE OF INVENTION: RECEPTORS FOR BOMBESIN-LIKE PEPTIDES
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Edwin P. Ching
STREET: 1501 Harbor Bay Parkway
CITY: Alameda
STATE: CA
COUNTRY: USA
ZIP: 94501
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/02091
FILING DATE: 19920313
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/426,150
FILING DATE: 24-OCT-1989
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/533,659
FILING DATE: 05-JUN-1990
ATTORNEY/AGENT INFORMATION:
NAME: Ching, Edwin P.
REGISTRATION NUMBER: 34090
REFERENCE/DOCKET NUMBER: A-0092C
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-266-7476
TELEFAX: 415-266-7400
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 384 amino acids
TYPE: AMINO ACID
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US92-02091-4

alignment_scores:

Quality: 466.00 Length: 311
 Ratio: 2.251 Gaps: 6
 Percent Similarity: 66.559 Percent Identity: 34.084

alignment_block:

US-09-931-157-2 x PCT-US92-02091-4 ..

Align seg 1/1 to: PCT-US92-02091-4 from: 1 to: 384

```

541 TACATCAACACGGTGTGTCCTGCTGCTGGTGGGATCATCGG 590
|||||
41 TyrValIleProAlaValTyrClyValIleIleLeuIleGlyLeuIleG1 57
|||||
591 GAACCTCCACACTTCGAGAATATATACAAGAACAAAGTCATGCCGAAACG 640
|||||
57 yAsnIleThrLeuIleLysIlePheCysThrValLysSerMetArgAsnV 74
|||||
641 GTCCCAATATCTTGATCCACGCTTGCTGCTGGGAGACCTGCTGCACATC 690
|||||
74 alProAsnLeuPheIleSerLeuAlaLeuGlyAspLeuLeuLeu 90
|||||
691 GTCATTCACATCCCTATCAATGTCTACAAAGCTGCTGGCAGAGGACTGCC 740
|||
91 IleThrCysAlaProValAspAlaSerArgTyrLeuAlaAspArgTrpLe 107
|||||
741 ATTTGGAGCTGAGATGTGAAGCTGGTGCCTTTTCATACAGAAAGCCTCCG 790
|||||
107 uPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLeuThrSerV 124
|||||
791 TGGGAATCACTGCTCAGTCTATGCTCTGCTGAGTATTGACAGATATCGA 840
|||||
124 alGlyValSerValPheThrLeuThAlaLeuSerAlaAspArgTyrLys 140
|||||
841 GCTGTGCTTCTTGGAGTAGAATTAAGGAATTTGGGTTCCAAATGGAC 890
|||
141 AlaIleValArgProMetAspIleGlnAlaSerHisAlaLeuMetLysI1 157
|||||
891 AGCAGTAGAATTTGTTGATTGGTGGTCTCTGTGGTCTGCTGGCTGCC 940
|||||
157 eCysLeuLysAlaAlaPheIleTrpIleIleSerMetLeuLeuAlaIleP 174
|||||
941 CTGAAGCCATAGTTTGTATATATTAATACGATGCACTACAAAGGAATTA 990
|||||
174 roGluAlaValPheSerAspLeuHisProPheHisGluSerThrAsn 190
|||||
991 CTGCGAATC.....TGCTTGTCTATCCGTTTCAGACACACCTTTCAT 1034
|||
191 GlnThrPheIleSerCysAlaProTyrProHisSerAsnGluLeuHisPr 207
|||||
1035 GCAGTTTTCACAGACAGCAAAAGATTGGTGGCTGTTTCAGTTTCTATTCT 1084
|||||
207 oLysIleHisSerMetAlaSer.....PheLeuValPheTyrV 220
|||||
1085 GCTTCCCATGGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
|||||
220 alIleProLeuSerIleIleSerValTyrTyrPheIleAlaLysAsn 236
|||||
1135 ATGTGTGAGAAAGAAAGTGGCAGATGCTTTAAATGATCACCTAAA 1184
|||||
237 LeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIleHisVally 253
|||||
1185 GCAG.....AGCGGGAAGTGGCCAAACCGTCTTTTGGCTGG 1222
|||
253 sLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeuValPheV 270
|||||
1223 TCCTTGTCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1272
|||||
270 alGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTyrLeuTyr 286
|||||
1273 AAGCTCACTCTTTATATATAGATATCCCAATAGATGTGAACCTTTGAG 1322
|||

```

```

287 ArgSerTyrHisTyrSerGluValAspThrSer.....MetLeuHi 300
1323 CTTTCTGTTGGTATTGGACTATATTGGTATCAACATGGCTTCACTGAATT 1372
|||
300 sPhe.....ValThrSerIleCysAlaArgLeuLeuAlaPheThrAsnS 315
1373 CCTGCATTAAACCAATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1422
|||||
315 erCysValAsnProPheAlaLeuTyrLeuLeuSerLysSerPheArgLys 331
1423 TGCCTTTAACTCATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1455
|||
332 GlnPheAsnThrGlnLeuLeuCys...CysGln 341

```

OM of: US-09-931-157-2 to: A_Geneseq_032802.* out_format : pfs

Date: Oct 1, 2002 1:04 AM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 CompuGen Ltd.

Command line parameters:

-MODEL=framer_n2p.model -DEV=xlp
-O=/cgn2.1/USPTO_spool/US09931157/runat_30092002_110420_11256/app_query.fasta_1.8657
-DB=A_Geneseq_032802 -OFMT=fastan -SUFFIX=rag -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-QGAPOP=4.500 -QGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09931157 -CGNL_1_398 -NCPU=6 -ICPU=3 -LONGLOG
-DEV=TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-931-157-2

Query length: 4301

Database: A_Geneseq_032802.*

Database sequences: 747574

Database length: 111073796

Search time (sec): 320.970000

score_list:

Sequence	Strd Orig	zScore	E Score	Len	Documentation
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: AAR30886		2340.00	4272.36	1.8e-230	44
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: AAB56352		2335.00	4263.16	5.9e-230	44
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: AAW59436		2325.00	4244.49	6.3e-229	45
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR25394		2044.50	3728.49	3.6e-200	44
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR43992		2034.50	3710.08	3.8e-199	44
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48720		1318.50	2394.95	8.5e-126	35
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAW02692		1318.50	2394.95	8.5e-126	35
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: AAR30885		1240.00	2248.10	1.1e-117	42
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR53078		1240.00	2248.10	1.1e-117	42
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB56351		1235.00	2238.90	3.4e-117	42
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48719		1166.00	2114.85	3.6e-110	33
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAW02691		1166.00	2114.85	3.6e-110	33
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR27507		506.50	899.28	1.6e-42	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: AAR12399		498.50	884.55	1.0e-41	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB66620		498.50	884.55	1.0e-41	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: AAR12125		496.50	880.87	1.7e-41	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR27510		494.00	876.08	3.1e-41	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY57956		491.00	870.56	6.2e-41	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB56369		490.00	868.72	7.9e-41	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB66621		489.50	867.80	8.8e-41	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR27509		486.50	862.28	1.8e-40	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR27508		467.00	826.57	1.8e-38	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB56359		461.00	815.53	7.3e-38	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48722		452.00	800.24	5.8e-37	34
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAW02694		452.00	800.24	5.8e-37	34
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAR89359		444.50	785.41	3.6e-36	37
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY58919		444.00	783.78	4.1e-36	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB47512		444.00	783.78	4.1e-36	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: AAR27511		443.00	781.94	5.2e-36	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAW07635		441.50	779.89	7.2e-36	37
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB56333		438.00	772.74	1.7e-35	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48721		436.50	772.31	2.2e-35	32
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAW02693		436.50	772.31	2.2e-35	32
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAB70002		434.50	766.81	3.8e-35	38
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: AAW07638		429.50	757.09	1.3e-34	39
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY90667		411.00	717.94	1.3e-32	61
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY90633		408.00	712.42	2.5e-32	61
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: AAW50303		398.50	696.40	2.2e-31	54
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY90665		387.00	678.49	2.5e-30	48
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: AAW50305		387.00	676.65	3.2e-30	48

/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: AAY32886 + 387.00 676.65 3.2e-3
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: AAY25969 + 387.00 676.65 3.2e-3
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: AAY90631 + 387.00 676.65 3.2e-3
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: AAR98261 + 385.50 671.01 5.1e-3
/SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: AAW37799 + 385.50 671.01 5.1e-3

seq_name: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: AAR30886

seq_documentation_block:

ID AAR30886 standard; Protein; 442 AA.

AC AAR30886;

XX 11-MAY-1993 (first entry)

DT Erb receptor.

DE Human; ETA; Erb; endothelin; receptor; transmembrane domain; N tail;
KW extracellular; cytoplasmic; C tail; post translational; bovine;
KW modification; ET-1 receptor; antagonist; circulatory system.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..26 /note= "Signal peptide"

FT Protein 27..442 /note= "Mature peptide"

FT Domain 102..126 /note= "Transmembrane domain I"

FT Domain 138..163 /note= "Transmembrane domain II"

FT Domain 176..197 /note= "Transmembrane domain III"

FT Domain 219..243 /note= "Transmembrane domain IV"

FT Domain 272..296 /note= "Transmembrane domain V"

FT Domain 325..350 /note= "Transmembrane domain VI"

FT Domain 363..389 /note= "Transmembrane domain VII"

FT Modified-site 59 /note= "N-glycosylated site"

FT Modified-site 305 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 413 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 419 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 435 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 436 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 440 /note= "May be phosphorylated with serine/threonine kinases"

XX EP522868-A.

PD 13-JAN-1993.

XX 92EP-0306347.

PF 10-JUL-1992;

PR 12-JUL-1991;

XX (SHIO) SHIONOGI SEIYAKU KK.

PI Imura H, Nakanishi S, Nakao K;
 XX WPI: 1993-010677/02.
 DR N-PSDB; AAQ34584.
 XX
 PT Human ETA and ETB endothelin receptors - for measuring endothelin
 PT and screening for endothelin antagonists
 XX
 PS Claim 8; Fig 2; 39pp; English.
 XX
 CC The sequences given in AAR30885-86 represent the human Eta and ETB
 CC endothelin receptors respectively. Eta is a 427 amino acid protein
 CC with a molecular weight of 48,726. ETB comprises 442 amino acids and
 CC has a molecular weight of 49,629. Eta has a higher affinity for
 CC endothelin (ET)-1 and ET-2, whereas ETB has no selectivity for ET-1,
 CC ET-2 or ET-3. The receptors each contain seven transmembrane domains
 CC and have an extracellular N tail and a cytoplasmic C tail. There are
 CC several potential sites for post translational modification, these
 CC sites are identical to those of bovine ET-1 receptor. Eta cDNA is
 CC 91.2% homologous to bovine ET-1 receptor cDNA and ETB cDNA is 61.1%
 CC homologous to that of bovine ET-1 receptor. The receptor proteins are
 CC useful as reagents for measuring the amount of ET or screening for
 CC antagonists of the ET receptor when studying the circulatory system.
 XX
 SO Sequence 442 AA;

alignment_scores:
 Quality: 2340.00 Length: 442
 Ratio: 5.294 Gaps: 0
 Percent Similarity: 100.000 Percent Identity: 100.000
 alignment_block:
 US-09-931-157-2 x AAR30886 ..
 Align seg 1/1 to: AAR30886 from: 1 to: 442

238 ATGCAGCGCGCTCCAAAGTCTGTGGGAGCGCGCTGTGGCTGGTCT 287
 1 MetGlnProProSerLeuLeuGlyArgAlaLeuValAlaLeuValLe 17
 288 TGCGCTGGCGCTGTGGGAGTCTGGGAGAGGAGAGAGGCTCCCGCTG 337
 17 uAlaCysGlyLeuSerArgIleTrpGlyGluArgGlyPheProA 34
 338 ACAGGCGCACTCCGCTTTTCCAAACCCAGAGATATGACGCCACCT 387
 34 spArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProThr 50
 388 AAGACCTTATGGCCCAAGGTTCCAGCGCAGTCTGGCGGCTGGTGGC 437
 51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeuAl 67
 438 ACCTGGGAGGTGCTTAAAGGACACAGGAGGAGGATCTCCCGCAGCA 487
 67 aProAlaGluValProLysGlySerArgThrAlaGlySerProArg 84
 488 CCATCTCCCTCCCGCTGCCAAGGAGCCATCGAGATCAAGGAGCTTC 537
 84 hrIleSerProProCysGlnGlyProIleGluIleLysGluThrPhe 100
 538 AAATACATCAACACGGTGTGCTGCTGCTGTGCTGGGATCAT 587
 101 LysThrIleAsnThrValValSerCysLeuValPheValLeuGlyIleI 117
 588 CGGGAATCCACACTCTCAGAAATATCTACAAGAACAGTGCATCGCAA 637
 117 eGlyAsnSerThrLeuLeuArgIleIleTrpLysAsnLysCysMetArg 134
 638 ACGGTCCCAATATCTTGATCGCCAGCTTGGCTCTGGGAGCTGTCGAC 687
 134 snGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHis 150

588 ATCGTCATTCATCCCTATCAATGCTTACAAGCTGTGGCAGAGGACTG 737
 151 IleValIleAspIleProIleAsnValTrpLysLeuLeuAlaGluAspTr 167
 738 GCCATTTGGAGCTGAGATGCTAAGCTGGTGCCTTTTCATACAGAAAGCT 787
 167 pProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
 788 CCCTGGGAATCACGTGCTAGTCTATGCTCTGAGTATTGACAGATAT 837
 184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTr 200
 838 CGAGCTTCTCTCTTGGAGTAGAATTAAGGAATTTGGGTTCCAAATG 887
 201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
 888 GACAGCAGTAGAATTTGTTGATTGGGTGGTCTCTGTGGTCTCTGCTG 937
 217 pThrAlaValGluIleValLeuIleTrpValValSerValValLeuAla 234
 938 TCCTGAAGCCATAGCTTTTGTATATATAGATGGACTACAAAGGAAGT 987
 234 aProGluAlaIleGlyPheAspIleIleThrMetAspTrpLysGlySer 250
 988 TATCTGCGAATCTGCTTGTCTTCATCCCTTCAGAAGACAGCTTTCATGCA 1037
 251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetG 267
 1038 GTTTTACAAGACAGCAAAAGATTGGTGGTGTTCAGTTTCTATTCTGCT 1087
 267 nPheTrpLysThrAlaLysAspTrpTrpLeuPheSerPheTrpPheCys 284
 1088 TGCCATTGGCCATCACTGCATTTTATATACACTATGACCTGTGAAATG 1137
 284 euProLeuAlaIleThrAlaPheTrpThrLeuMetThrCysGluMet 300
 1138 TTGAGAAAGAAAGTGGCATGCAGATGCTTTAAATGATCACTAAAGCA 1187
 301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
 1188 GAGAGGGAAGTGGCCAAACCGCTTTTGGCTGCTCTGTGCTTTGGCC 1237
 317 nArgGluValAlaLysThrValPheCysLeuValLeuValPheAla 334
 1238 TCTGCTGGCTTCCCTTCACCTCAGCAGATTCGAACTCACTCTTTAT 1287
 334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTr 350
 1288 AATCAGATGATCCCAATAGATGTGAACTTTTGAGCTTCTGTGGTATT 1337
 351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuVal 367
 1338 GGACTATTTGGTATCAACATGGCTTCACTGAATTCCTGCAATTAACCAA 1387
 367 uAspTrpIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnPro 384
 1388 TTGCTCTGATTTTGGTGAGCAAAAGATTCAAAACTGCTTTAAGTCATGC 1437
 384 lealaLeuTrpLeuValSerLysArgPheLysAsnCysPheLysSerCys 400
 1438 TTATGCTGTGGTGGCAGTCAATTTGAGAAAAACAGTCTTGGAGGAAA 1487
 401 LeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluGlu 417
 1488 CGAGTCTGCTTAAAGTTCAAGCTAATGATCAGGATATGACAATTC 1537
 417 sGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTrpAspAsnPhe 434
 1538 GTTCCAGTAATAAATACAGCTCATCT 1563
 434 rGserSerAsnLysTrpSerSer 442

seq_name: /SID1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.AB56352

```
seq_documentation_block:
ID   ABB56352 standard; Protein; 442 AA.
XX   AC   ABB56352;
XX   DT   18-FEB-2002 (first entry)
XX   DE   Non-endogenous human GPCR protein, SEQ ID NO: 497.
XX   KW   Human; G protein-coupled receptor; GPCR; non-endogenous; mutant;
XX   KW   constitutively activated GPCR; agonist; disease.
XX   OS   Homo sapiens.
XX   OS   Synthetic.
XX   PN   WO200177172-A2.
XX   PD   18-OCT-2001.
XX   PF   05-APR-2001; 2001WO-US11098.
XX   PR   07-APR-2000; 2000US-195747P.
XX   PA   (AREN-) ARENA PHARM INC.
XX   PI   Lehmann-Bruinsma K, Liaw CW, Lin I;
XX   DR   WPI; 2001-648759/74.
XX   DR   N-PSDB; ABI97988.
XX   PT   Identifying agonists of G protein-coupled receptors (GPCRs) for use in
XX   PT   disease treatment, comprises contacting candidate compounds with
XX   PT   versions of GPCRs.
XX   PS   Claim 1; Page 297-299; 394pp; English.
XX   CC   The invention relates to G protein-coupled receptors (GPCRs) for which
XX   CC   the endogenous ligand has been identified. Non-endogenous
XX   CC   constitutively activated versions of known GPCRs are used
XX   CC   in the invention for the direct identification of candidate compounds as
XX   CC   receptor agonists, inverse agonists or partial agonists. Such
XX   CC   agonists are useful as therapeutic agents for diseases or disorders
XX   CC   associated with GPCRs. The present sequence is a non-endogenous
XX   CC   version of a known human GPCR.
XX   SQ   Sequence 442 AA;

alignment_scores:
Quality: 2335.00      Length: 442
Ratio: 5.295          Gaps: 0
Percent Similarity: 99.774 Percent Identity: 99.774

alignment_block:
US-09-931-157-2 x ABB56352
Align seg 1/1 to: ABB56352 from: 1 to: 442

238 ATCCAGCGCGCTCCAACTGTGTCGGACGGCGCTGGTTCGGTGGTCT 287
1 MetGlnProProSerLeuCysGlyArgAlaValAlaLeuValle 17
288 TGCGCTCGGCTGTCGGATCTGGGAGAGAGAGAGGCTTCCCGCGCTG 337
17 uAlaCysGlyLeuSerArgIleIrrpGlyGluGluArgGlyPheProPro 34
338 ACAGGCGCCACTCGCGCTTTTGAACCGCAGAGATAATGACGCCACCACT 387
34 spArgAlaThrProLeuLeuGlnThrAlaGluIleMetThrProProThr 50
388 AAGACCTTATGGCCCAAGGGTTCCAAACGGCAGTCTCGCGCGGTCTGGC 437
|||||
```

```
51 LysThrLeuTrpProLysGlySerAsnAlaSerLeuAlaArgSerLeuAl 67
438 ACCTGCGGAGGTGCTAAAGGACAGCAGCGGAGGATCTCCGCCACGCA 487
|||||
67 aProAlaGluValProLysGlyAspArgThrAlaGlySerProProArgT 84
488 CCATCTCCCTCCCGTCCCAAGAGCCCATCGAGATCAAGAGAGACTTTC 537
|||||
84 hrIleSerProProCysGlnGlyProIleGluIleLysGluThrPhe 100
538 AATACATCAACACGGTGTGCTCCTCGCTTGTGCTGCTGGGATCAT 587
|||||
101 LysTyrIleAsnThrValValSerCysLeuValPheValLeuGlyIleI 117
588 CGGAACTCCACACTTCTGAGAAATATCTACAAGAACAAGTGCATCGAA 637
|||||
117 eGlyAsnSerThrLeuLeuArgIleIleTyrLysAsnLysCysMetArg 134
638 ACGGTCCCAATATCTTGATCGCCAGCTTGCTCTGGGAGACCTGCTGCAC 687
|||||
134 snGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyAspLeuLeuHis 150
688 ATCGTCAATTGACATCCCTATCAATGCTCTACAAGCTGCTGCCAGAGACTG 737
|||||
151 IleValIleAspIleProIleAsnValTyrLysLeuLeuAlaGluAspTr 167
738 GCATTTGGAGCTGAGATGTGAAGCTGGTGCCTTTTCATACAGAAAGCCT 787
|||||
167 pProPheGlyAlaGluMetCysLysLeuValProPheIleGlnLysAla 184
788 CCGTGGGAATCACTGCTGAGTCTATGTGCTGCTGATATTGACAGATAT 837
|||||
184 erValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyr 200
838 CGAGCTGTTCCTTCTGGAGTAGAATTAAGGAATTGGGCTTCCAAATG 887
|||||
201 ArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValProLysTr 217
888 GACAGCAGTAGAAATTTGTTGATTTGGGTGGTCTCTGTGTTCTGCTGCTG 937
|||||
217 pThrAlaValGluIleValleuIleTrpValValSerValValLeuAlaV 234
938 TCCTCGAAGCCATAGTGTGTTGATATATACGATGAGTACAAAGGAAGT 987
|||||
234 alProGluAlaIleGlyPheAspIleIleThrMetAspTyrLysGlySer 250.
988 TATCTCGGAATCTGCTTGCCTTCATCCCGTTCAGAACACACACCTTTCATGCA 1037
|||||
251 TyrLeuArgIleCysLeuLeuHisProValGlnLysThrAlaPheMetG 267
1038 GTTTTACAAGACAGCAAAAGATTGGTGGCTGTTTCAGTTCATTCTATTCTGCT 1087
|||||
267 nPheTyrLysThrAlaLysAspTrpIlePheSerPheTyrPheCysL 284
1088 TGCCATTGGCCATCACTGCATTTTATATACACTAATGACCTGTGAATG 1137
|||||
284 euProLeuAlaIleThrAlaPhePheTyrThrLeuMetThrCysGluMet 300
1138 TTGAGAAAAGAAAGTGGCATGCGATGCTTTTAAATGATCACCCTAAAGCA 1187
|||||
301 LeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysG 317
1188 GAGACGGGAAGTGGCCAAACCGTCTTTTCCCTGGTCTGCTGCTTTGGCC 1237
|||||
317 nArgArgGluValLysLysThrValPheCysLeuValLeuValPheAlaL 334
1238 TCTGCTGGCTTCCCTTCACTCAGCAGGATTCCTGAAGCTCCTCTCTTAT 1287
|||||
334 euCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyr 350
1288 AATCAGAAATGATCCCAATAGATGTGAACCTTTTGAGCTTCTCTGTGTATT 1337
|||||
351 AsnGlnAsnAspProAsnArgCysGluLeuLeuSerPheLeuLeuValle 367
```



```

435 GGCACCTGGGAGGTCCTAAAGGAGACAGGAGCGGAGGATCCGCGCAC 484
66 eargThrAlaGluValThrLysGlyGlyArgValAlaGlyValProA 83
485 GCACCATCCCTCCCTCCCGGCCAAGGACCATCGAGATCAAGGAGACT 534
83 rgSerPhe...ProProCysGlnArgLysIleGluIleAsnLysThr 98
535 TTCAAATACATCAACACAGGTTGTCTCCCTGCTGCTGCTGCTGGGAT 584
99 PheLysThrIleAsnThrIleValSerCysLeuValPheValLeuGlyI 115
585 CATCGGGAATCCACACTCTCGAAGATATCTACAAGAACAGTGCATGC 634
115 eileGlyAsnSerThrLeuLeuArgIleIleThrLysAsnLysCysMe 132
635 GAAACGGTCCCAATATCTGATCCCGAGCTGGCTGCGGAGAGCTGCTG 684
132 rgAsnGlyProAsnIleLeuIleAsrLeuAlaLeuGlyAsPheLeu 148
685 CACATCTCATTTGACATCCCTATCAATCTTACAAGCTGCTGGCAGAG 734
149 HisIleIleIleAspIleProIleAsnAlaThrLysLeuLeuAlaG 165
735 CTGCCATTTGGAGCTGAGATGTGTAGCTGGTGGCTTTCATACAGAA 784
165 pTrpProPheGlyAlaGluMetCysLysLeuValProPheIleGln 182
785 CPTCCGTGGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 834
182 lAsrValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAsp 198
835 TATCGAGCTGTTGCTTCTTGGAGTAGAATTAAGGAATTTGGGTTCC 884
199 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleGlyValP 215
885 ATGGACACAGTAGAATTTGTTGATTTGGTGGTCTCTGTTGTTCTG 934
215 strPthrAlaValGluIleValLeuIleTrpValValSerValVal 232
935 CTGCTCCCTGAAGCCATAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 984
232 lavalProGluAlaIleGlyPheAspValIleThrSerAspThrLys 248
985 AGTATCTGCGAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1034
249 LysProLeuArgValCysMetLeuAsnProPheGlnLysThrAlaP 265
1035 GCAGTTTACAGACAGCAAAAGATTTGCTGCTGCTGCTGCTGCTGCT 1084
265 tGlnPheThrLysThrAlaLysAspTrpLeuPheSerPheThrPhe 282
1085 GCTTGCATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1134
282 ysLeuProLeuAlaIleThrAlaIlePheThrLeuMetThrCysG 298
1135 ATGTTGAGAAGAAAGTGGCATCGAGATTCCTTAAATGATCACTAAA 1184
299 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHis 315
1185 GCAGAGACGGAGAGTGGCCAAACCGTCTTTGCTGCTGCTGCTGCT 1234
315 sGlnArgGluValAlaLysThrValPheCysLeuValLeuValPhe 332
1235 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1284
332 lalLeuCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuTh 348
1285 TATAATCAGATATCCCATATGATGTAATTTGAGCTTTTGGCTTGT 1334
349 TyrAspGlnSerAsnProGlnArgCysGluLeuLeuSerPheLeu 365
1335 ATTGGACTATATTGGTATCAACATGCTTCACTGCTGCTGCTGCTG 1384

```

```

|||||
365 lLeuAspTyrIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsn 382
1385 CAATTGCTGCTGATTTGGTGAGCAAAAGATTCAAAAGCTGTTTAAAGTCA 1434
|||||
382 tollAlaLeuThrIleValSerLysArgPheLysAsnSerCysPheLys 398
1435 TGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1484
399 CysLeuCysCysTrpCysGlnThrPheGluGluLysGlnSerLeuIu 415
1485 AAACGAGCTGCTGCTTAAAGTTCAAAAGCTGATGATGATGATGATG 1534
415 uLysGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyrAsp 432
1535 TCCGTTCCAGTAATAATACAGCTCATCT 1563
|||||
432 heArgSerSerAsnLysTyrSerSer 441
seq_name: /SIDS1/gcgdata/geneseq/geneseq-embl/AA1994.DAT:AA43992
seq_documentation_block:
ID AAR43992 standard; Protein; 441 AA.
XX
AC AAR43992;
XX
DT 28-JUN-1994 (first entry)
XX
DE Bovine ET receptor subtype 2.
XX
KW Bovine; endothelin; ET; receptor; subtype 2; prevention; lung;
KW ischaemic tissue interference; diagnosis.
XX
OS Bos taurus.
XX
FH Key
FT Peptide
FT Protein
FT Protein
FT Protein
XX
XX JP05310793-A.
XX
XX 22-NOV-1993.
XX
XX 21-NOV-1991; 91JP-0306429.
XX
XX 05-SEP-1991; 91JP-0226204.
XX
XX (CHCH ) CHICHIBU CEMENT KK.
XX
XX WPI; 1994-002317/01.
XX
XX N-PSDB; AAQ53922.
XX
XX Bovine endothelin receptor sub-type B - used for elucidation of
XX mechanism of ischaemic tissue interference and its treatment,
XX diagnosis and prevention
XX
XX Claim 1; Fig 1; 9pp; Japanese.
XX
XX This sequence represents the bovine endothelin (ET) receptor subtype
XX 2. This receptor protein may be used for the elucidation of the
XX mechanism of ischaemic tissue interference and its treatment,
XX diagnosis and prevention. The cDNA encoding this sequence was
XX isolated from bovine lung tissue.
XX
XX Sequence 441 AA;

```

```

alignment_scores:
Quality: 2034.50
Ratio: 4.902
Percent Similarity: 93.679
Length: 443
Gaps: 2
Percent Identity: 88.036

```

alignment_block:

US-09-931-157-2 x AAR43992 ..

Align seg 1/1 to: AAR43992 from: 1 to: 441

```

238 ATGCACCGCGCTCCAGTCTGTGGACGCGCGCTGTGTGGCTGTTCT 287
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
1 MetGlnProLeuProSerLeuCysGlyArgAlaLeuValAlaLeuIleLe 17
288 TGCGCTCGCGCTGCTCCGGATCTGGGGAGAGAGAGAGGCTTCCCGCGCTG 337
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
17 uAlaCysGlyValAlaGlyIleGlnAlaGluArgGluPheProProA 34
338 ACAGGGCCACT...CGCGTTTTCACAAACGCGAGAGATATGACGCCACCC 384
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
34 laGlyAlaThrGlnProLeuProGlyThrGlyGluMetMetGluThrPro 50
385 ACTAAGACCTTATGGCCCAAGGTTCCAGCGGCAGCTCTGGCGGCGT 434
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
51 ThrGluThrSerTrpProGlyArgSerAsnAlaSerAspProArgSerSe 67
435 GGCACCTGGCGAGTGCCTAAAGACAGACAGGACGCGAGGATCTCCGCCAC 484
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
67 rAlaThrProGlnIleProArgGlyArgMetAlaGlyIleProProA 84
485 GCACCATCTCCCTCCCGTGCACAGGCCCATCGAGATCAAGAGACT 534
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
84 rgThr.....ProProCysAspGlyProIleGluIleLysGluThr 98
535 TTCAATATACATCAACACGGTTGTGCTCCTGCTGTGTTGCTGCTGGGAT 584
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
99 PheLysTyriLeAsnThrValValSerCysLeuValPheValLeuGlyI 115
585 CATCGGGAACCTCCACACTTCTGAGAATTATCTACAAGAACAACTGCATGC 634
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
115 elleGlyAsnSerThrLeuLeuArgIleIleTyriLysAsnLysCysMeta 132
635 GAAACGGTCCCAATATCTTGCATCGCCAGCTTGCTGCGAGACCTCGCTG 684
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
132 rgAsnGlyProAsnIleLeuIleAlaSerLeuAlaLeuGlyValLeuLeu 148
685 CACATCGCTATTGACATCCCTATCATGTCTACAAGCTGCTGCGCAGGA 734
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
149 HisIleIleIleAspIleProIleAsnThrTyriLysLeuLeuAlaLysAs 165
735 CTGGCCATTGGAGCTGAGATGTGTAAGCTGGTGCCTTTCATACAGAAAG 784
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
165 pTrpProPheGlyValGluMetCysLysLeuValProPheIleGlnLysA 182
785 CTCTCGTGGGAATCACTGTGCTGAGTCTATGTGCTCTGAGTATGACAGA 834
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
182 laSerValGlyIleThrValLeuSerLeuCysAlaLeuSerIleAspArg 198
835 TATCGAGCTGTCTGCTTCTGGAGTACAATTAAAGGAATTGGGGTCCAAA 884
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
199 TyrArgAlaValAlaSerTrpSerArgIleLysGlyIleCylValProLy 215
885 ATGCACAGCAGTAGAATAATGTTTGTATTTGGTGGTCTCTGTGTTCTGG 934
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
215 sTrpThrAlaValGluIleLeuLeuIleTrpValValSerValValLeuA 232
935 CTCTCCCTGAAGCCATAGGTTTGTATATATTACGATGGACTACAAGGA 984
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
232 laValProGluAlaValGlyPheAspIleIleThrSerAspHisIleGly 248
985 AGTTATCTCGGATCTGCTGCTGCTTCATCCCGTTCACAGACACACCTTCAT 1034
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
249 AsnLysLeuArgIleCysLeuLeuHisProThrGlnLysThrAlaPheMe 265
1035 GCAGTTTTCACACAGCAGCAAGATTGGTGGCTGTTCAAGTTCATTCTTCT 1084
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
265 tGlnPheTyriLysThrAlaLysAspTrpTrpLeuPheSerPheThrPheC 282

```

```

1085 GCTTGCCATTGGCCATCACTCACTATTTTATACACTAATGACCTGTGAA 1134
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
282 ysLeuProLeuAlaIleThrAlaLeuPheTyriThrLeuMetThrCysGlu 298
1135 ATGTTGAGAAAGAAAGTGGCATGCGATTGCTTTAAATGATCACCTAAA 1184
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
299 MetLeuArgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLy 315
1185 GCAGACACGGAGTGGCCAAACCCCTTTTCCCTGGCTGCTTGTCTTG 1234
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
315 sGlnArgArgGluValAlaLysThrValPheCysLeuValLeuValPheA 332
1235 CCCTCTGCTGGCTTCCCTTCACTCAGCAGGATTCTGAAGCTCACTCTT 1284
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
332 laLeucCysTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeu 348
1285 TATAATCAGAAATGCCAATAGATGTGAACCTTTTGTAGCTTTCTGTGGT 1334
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
349 TyrAspGlnHisAspProArgCysGluPheLeuSerPheLeuLeuVa 365
1335 ATTGGACTATATGGTATCAACATGGCTTCACTGAATTCCTGCTAATACC 1384
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
365 lleAspTyriIleGlyIleAsnMetAlaSerLeuAsnSerCysIleAsnP 382
1385 CAATTCTCTGTATTGTTGGTGGACAAAGATTCAAAACACTGCTTTAAGTCA 1434
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
382 roIleAlaLeuTyriLeuValSerLysArgPheLysAsnCysPheLysSer 398
1435 TGCTTATGCTGCTGGTGGCAGTCATTTGAAGAAAAACAGCTCCTTGGAGGA 1484
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
399 CysLeuCysCysTrpCysGlnSerPheGluGluLysGlnSerLeuGluG 415
1485 AAAGCAGTCTGCTTAAAGTTCAAAGTCAATGATCAGGATATGACAACT 1534
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
415 ulyGlnSerCysLeuLysPheLysAlaAsnAspHisGlyTyriAspAsnP 432
1535 TCGTTTCCAGTAAATAACAGCTCATCT 1563
|||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||  |||||||
432 heArgSerSerAsnLysTyriSerSerSer 441

```

seq_name: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: AAR48720

seq_documentation_block:

ID AAR48720 standard; Protein; 350 AA.

XX AAR48720;

XX 06-JUN-1996 (first entry)

XX G-protein coupled rat non-isopeptide selective endothelin receptor.

XX G-protein coupled receptor; ligand binding assay; transmembrane domain;
 KW psychotic disorder; schizophrenia; dopamine; cAMP; adenosine; thrombin;
 KW muscarinic acetylcholine; adrenergic; endothelin; bombesin; endocrine;
 KW rhodopsin; opsin; odorant; cytomagalovirus.

XX Rattus rattus.

XX W09405695-A1.

XX 17-MAR-1994.

XX 09-SEP-1993; 93WO-US08528.

XX 10-SEP-1992; 92US-0943236.

XX (UJNY) UNIV NEW YORK STATE.

XX Murphy RB, Schuster DI;

XX WPI; 1994-101120/12.

XX

PT Polypeptides of G-coupled receptor proteins (GPRs) - useful for
PT binding GPR ligands or modulating GPR binding
xy

PS
yy Disclosure; Page 95-96; 160pp; English.

Proteins AAR48685-R48758 represent a range of G-protein coupled receptor proteins selected from cAMP, adenosine, muscarinic acetylcholine, adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin, odorant, cytomegaloviral and other G-protein coupled receptors. The receptor proteins were used to design polypeptides, pref. based on the transmembrane domains, for use in G-protein coupled receptor ligand binding assays. The polypeptide fragments retain biological activity such as binding a GPR ligand or modulating GPR ligand binding to a GPR (see AAR48759-R48758, AAR50569-R50807 and AAR89189-R89195 for examples of polypeptide fragments). The polypeptide fragments can be used in compositions for treating subjects suffering from a pathology related to a GPR abnormality e.g. a psychotic disorder such as schizophrenia.

Sequence 350 AA:

alignment_scores:			
Quality:	1318.50	Length:	274
Ratio:	4.938	Gaps:	3
Percent Similarity:	97.445	Percent Identity:	91.606

alignment_block:

US-09-931-157-2 x AAR48720

Align seg 1/1 to: AAR48720 from: 1 to: 350

745 GGAGCTGAGATGTGTAAGCTGTGGTCCCTTCATACAGAAAGCCTCCGTGG 794
 |||CysArgLeuCysArgLeuValProPheIleGlnLysAlaSerValG1 95
 795 AATCACTGTGCTCAGCTCTATGTGCTCTGAGTATTGACAGATATCGAGCTG 844
 |||ThrValLeuSerLeuCysAlaLeuSerIleAspArgTyrArgAlaVal 112
 845 TTGCTCTCTGGAGTAGAATAAAGGAATTGGGTGCCCAAAATGGACAGCA 894
 |||AlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpThrAla 128
 112 alAserTrpSerArgIleLysGlyIleGlyValProLysTrpThrAla 128
 895 GTAGAAATGTTTGGTGGGTGCTCTCTGTGGTCTGTGGTGCCTCGCA 944
 |||ValGluIleValLeuIleTrpValValSerValValLeuAlaValProG1 145
 129 ValGluIleValLeuIleTrpValValSerValValLeuAlaValProG1 145
 945 AGCCATAGTGTGATATTAATACATCGACTACAAAGGAAGTATATCGC 994
 |||unalleGlyPheAsp...ThrThrSerAspTyrLysGlyLysProLeuA 161
 145 unalleGlyPheAsp...ThrThrSerAspTyrLysGlyLysProLeuA 161
 995 GAATCTGCTTGCCTTCATCCGTTACAGAGACAGCTTTCATCCAGTTTAC 1044
 |||ValCysMetLeuAsnProPheGlnLysThrAlaPheMet...PheTyr 176
 161 rgValCysMetLeuAsnProPheGlnLysThrAlaPheMet...PheTyr 176
 1045 AAGACA...GCAAAAGATGTGGCTGTTCAGTTTCTATTCTGCTGCTGCC 1091
 |||LysThrAlaAlaLysAspTrpTrpLeuPheAlaPheTyrPheCysLeuPr 193
 177 LysThrAlaAlaLysAspTrpTrpLeuPheAlaPheTyrPheCysLeuPr 193
 1092 ATTGGCCATCACTGCATTTTATACACTAATGACCTGTGAAATGTTGA 1141
 |||LeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 210
 193 oLeuAlaIleThrAlaIlePheTyrThrLeuMetThrCysGluMetLeuA 210
 1142 GAAAGAAAGTGGCATGCAGATTGCTTTAAATGATACACCTAAAGCAGAGA 1191
 |||rgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArg 226
 210 rgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArg 226
 1192 CGGGAAGTGGCCAAACCGTCTTTTGCTGGTCTCTGTCTTTGCCCTCTG 1241
 |||ArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaLeuCy 243
 227 ArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaLeuCy 243
 1242 CTGGCTTCCCTTCACCTCAGCAGGATCTGAAGCTCACTCTTATTAATC 1291
 |||

243 sTrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyrAspG 260

1292 AGAATGATCCCAATAGATGTGAACATTGTGAGCTTCTGTTGGTATTGGAC 1341

260 InSerAsnProGlnArqCysGluLeuLeuSerPheLeuLeuValLeuAsp 276

1342 TATATTGGTATCAACATGGCTTCACCTGAATTCCTGCATTAAACCCAATTGC 1391

277 TyrIleGlyLeuAsnMetalSerIleAsnSerCysIleAsnProIleAl 293

1392 TCTGTATTTCGGTCAGCAAAAGATTCAAAACTGCTTTAAAGTCATGCTTAT 1441

293 aLeuTyrLeuValSerLysArgPheLysAsnCysPheLysSerCysLeuC 310

1442 GTCTGCTGGTCGACGTCATTGTGAAGAAACAGTCCCTTCGAGGAACGAG 1491

310 yscYstrpcysGlnthrPheGluGluLysGlnSerLeuGluGluLysGln 326

1492 TCGTGCTTAAAGTTCAAAGCTAATGATCACGGATATGACAACTTCGGTTC 1541

327 SerCysLeuLysPheLysAlaAsnAspHisGlyTyrAspAsnPheArgSe 343

1542 CAGTAATAAATCACGCTCATCT 1563

343 rSerAsnLysTyrSerSerSer 350

seq_name: /SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:AAW02692

seq_documentation_block:

ID AAW02692 standard; peptide; 350 AA.

AAW02692;

XX
DT 12-NOV-1996 (first entry)

xx
DE
xx
xx

G-protein coupled receptor; ligand binding assay; transmembrane domain; schizophrenia; dopamine; AMP; adenosine; thrombin; adrenergic; opsin; muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin; odorant; cytomegalavirus; serotoninergic.

XX Rattus rattus.
OSXX
PN
US5508384-A.

XX
PD 16-APR-1996.

10-SEP-1992:

XX
PR
09-SEP-1993:
93US-0118270

PR 10-SEP-1992; 92US-0943236.
XX

PA (UUNY) UNIV NEW YORK STATE.
XX

PI
Murphy RB, Schuster DI;

XX
DR WPI; 1996-208785/21.

New dopamine receptor peptide - useful as antipsychotic agent, e.g. for treating schizophrenia

PS
Disclosure: Column 109-112; 184pp; English.

Proteins AA02657-W02730 represent a range of G-protein coupled receptor (GPR) proteins selected from cAMP, adenosine, muscarinic acetylcholine, adrenergic, thrombin, endothelin, bombesin, endocrine, rhodopsin, opsin, odorant, cytomegaloviral and other GPR proteins. The receptor proteins were used to design polypeptides, pref. based on the transmembrane domains, for use in G-protein coupled receptor ligand binding assays. The polypeptide fragments retain biological activity such as binding a GPR ligand or modulating GPR ligand binding to a GPR (see AA02747-W02999 for examples of polypeptide fragments). The polypeptide

CC fragments can be used in compositions for treating subjects suffering
CC from a pathology related to a GPR abnormality e.g. a psychotic disorder
XX such as schizophrenia.

SQ Sequence 350 AA:

alignment_scores:

Quality: 1318.50 Length: 274
Ratio: 4.938 Gaps: 3
Percent Similarity: 97.445 Percent Identity: 91.606

alignment_block:

US-09-931-157-2 x AA02692

Align seg 1/1 to: AA02692 from: 1 to: 350

745 GGAGCTGAGATGCTGAAGTGGTGGCTTCATACAGAAAGCCCTCGTGG 794
|||||
79 GlyCysArgLeuCysArgLeuValProPheIleGlnLysAlaSerValGI 95
795 AATCACTGTGCTGAGTCTATGCTCTGAGTATTGACAGATATCGAGCTG 844
|||||
95 ylleThrValLeuSerLeuCysAlaLeuSerIleAspArgTyArgAlav 112
845 TTGCTTCTTGGAGTAAAGAAATGGGGTTCCAAAATGGACAGCA 894
|||||
112 alAlaSerTrpSerArgIleLysGlyIleGlyValProLysTrpThrAla 128
895 GTAGAATTTGTTGATTTGGTGGTCTCTGTTGCTGCTGCTCCCTGA 944
|||||
129 ValGluIleValLeuIleTrpValValSerValValLeuAlaValProGI 145
945 AGCATAGTGTGATATATTAATACGATGAGTACAAAGGAAGTTATCTGC 994
|||||
145 uAlaIleGlyPheAsp...ThrThrSerAspTyLysGlyLysProLeuA 161
995 GAATCTGCTTGGTTCATCCCGTTCAGAGACAGCTTTCATGAGTTTAC 1044
|||||
161 rgValCysMetLeuAsnProPheGlnLysThrAlaPheMet...PheTyr 176
1045 AAGACA...GCAAAAGATGGTGGCTTCAGTTTCTTCTATTTCTGCTGCC 1091
|||||
177 LysThrAlaAlaLysAspTrpIlePheAlaPheTyLysPheCysLeuPr 193
1092 ATGGCCATCACTGCTATTTTATACATTAATGACTGTGAAATGTTGA 1141
|||||
193 OLeuAlaIleThrAlaIlePheTyThrLeuMetThrCysGluMetLeuA 210
1142 GAAAGAAAGTGCGATGCGAGATGCTTTAAATGATCACCTAAAGCAGAGA 1191
|||||
210 rgLysLysSerGlyMetGlnIleAlaLeuAsnAspHisLeuLysGlnArg 226
1192 CGGGAAGTGGCCAAACCGCTTTTGGCTGGTCTGCTTTGCCCTCTG 1241
|||||
227 ArgGluValAlaLysThrValPheCysLeuValLeuValPheAlaLeuCy 243
1242 CTGGCTTCCCTTCACCTCAGCAGGATTTCTGAAGCTCACTCTTATAATC 1291
|||||
243 StrpLeuProLeuHisLeuSerArgIleLeuLysLeuThrLeuTyAspG 260
1292 AGAATGATCCCAATAGATGTAACCTTTTGAGCTTTCTGTGTTGTTGGAC 1341
|||||
260 InSerAsnProGlnArgCysGluLeuLeuSerPheLeuLeuValLeuAsp 276
1342 TATATTGGTATCAACATGCTTCACCTGAATTCCTGCAATTAACCAATTCG 1391
|||||
277 TyrIleGlyIleAsnMetAlaSerIleAsnSerCysIleAsnProIleAl 293
1392 TCTGTATTGGTGAGCAAAAGATTCAAAACCTTTTAAGTCATGCTTAT 1441
|||||
293 AleuTyLeuValSerLysArgPheLysAsnCysPheLysSerCysLeuC 310

1442 GCTGCTGGTGGCCAGTCATTGTAAGAAAAACAGTCCTTGGAGGAAAGCAG 1491
|||||
310 yscysTrpCysGlnThrPheGluGluLysGlnSerLeuGluGluLysGln 326
1492 TCGTGCTTAAAGTTCAAGCTATATGATCAGCGATATGACAACTTCCTTC 1541
|||||
327 SerCysLeuLysPheLysAlaAsnAspHisGlyTyAspAsnPheArgse 343
1542 CAGTAATAAATACAGCTCATCT 1563
|||||
343 rSerAsnLysTySerSerSer 350

seq_name: /SIDS1/gc9data/geneseq/geneseq-emb1/AA1993.DAT:AA030885

seq_documentation_block:

ID AAR30885 standard; Protein; 427 AA.

AC AAR30885;

DT 11-MAY-1993 (first entry)

XX Eta receptor.

DE Human; Eta; ETb; endothelin; receptor; transmembrane domain; N tail;
KW extracellular; cytoplasmic; C tail; post translational; bovine;
KW modification; ET-1 receptor; antagonist; circulatory system.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..20 /note= "Signal peptide"

FT Protein 21..427 /note= "Mature peptide"

FT Domain 81..105 /note= "Transmembrane domain I"

FT Domain 117..139 /note= "Transmembrane domain II"

FT Domain 160..181 /note= "Transmembrane domain III"

FT Domain 204..229 /note= "Transmembrane domain IV"

FT Domain 257..278 /note= "Transmembrane domain V"

FT Domain 308..332 /note= "Transmembrane domain VI"

FT Domain 347..372 /note= "Transmembrane domain VII"

FT Modified-site 29 /note= "N-glycosylated site"

FT Modified-site 62 /note= "N-glycosylated site"

FT Modified-site 379 /note= "May be palmitoylated as in beta-2-adrenergic receptor"

FT Modified-site 383 /note= "May be palmitoylated as in beta-2-adrenergic receptor"

FT Modified-site 385..388 /note= "One base may be palmitoylated as in beta-2-adrenergic receptor"

FT Modified-site 289 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 391 /note= "May be phosphorylated with serine/threonine kinases"

FT Modified-site 421 /note= "May be phosphorylated with serine/threonine kinases"

PN EP522868-A.

XX


```

seq_name: /SIN1/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:AA848719
seq_documentation_block:
ID   AAR48719 standard; Protein; 333 AA.
XX
AC   AAR48719;
XX
XX
DT   06-JUN-1996 (first entry)
XX
DE   G-protein coupled cow endothelin 1 receptor protein.
XX

```



```

seq_documentation_block:
ID   AAR12399 standard; Protein; 384 AA.
XX
AC   AAR12399;
XX
DT   01-AUG-1991 (first entry)
XX
DE   Gastrin releasing peptide receptor.
XX
KW   GRPR.
XX
OS   Mus musculus.
XX
PN   W09106647-A.
XX
PD   16-MAY-1991.
XX
PF   23-OCT-1990; 90WO-US06125.
XX
PR   23-OCT-1990; 90WO-US06125.
XX
PR   24-OCT-1989; 89US-0426150.
XX
PR   05-JUN-1990; 90US-0533659.
XX
PA   (TRIT-) TRITON BIOSCIENCES.
XX
PI   Feldman RI, Harkins RN, Batey J, Slattey TK;
XX
DR   WPI; 1991-164196/22.
XX
DR   N-PSDB; AAQ12032.
XX
PT   DNA and polypeptide sequence of GRP receptor - used in treatment
XX
PT   and diagnosis of cancerous tissues.
XX
PS   Claim 5; Fig 8; 84pp; English.
XX
CC   The purified receptor can be used in drug screening for cpds.
XX
CC   having suitable binding affinity for GRPR. The receptor,
XX
CC   fragments and antibodies may also be used in diagnostic assays
XX
CC   to determine altered levels of GRPR. The antibodies are pref.
XX
CC   raised to fragments comprising residues 1-39, 64-77, 98-115,
XX
CC   138-157, 176-209, 236-266, 288-300 and 330-385.
XX
SQ   Sequence 384 AA;

alignment_scores:
    Quality: 498.50          Length: 362
    Ratio: 2.149             Gaps: 8
Percent Similarity: 64.088 Percent Identity: 33.702

alignment_block:
US-09-931-157-2 x AAR12399 ..

Align seg 1/1 to: AAR12399 from: 1 to: 384

484 CGCACCATCTCCCTCCCGTCGCCAGGACCCATCGATCAAGAGAC 533
      :::::::::::::::::::: :
25 GlnSerLeuSerProProLysMetAspAsn.....TrpPheHisProGl 39
      yffllyrrvVallleProAlValTyrgLyLeullelleValIleGlyL 56
584 TCATCGGGAACCTCACACTTCTTGAGAATTATTCACAAGAACAAAGTCATG 633
      :::::::::::::::::::: :
56 euileGlyAsnIleThrLeuilleLysillePheCysThrValIysSerMet 72.
634 CGAAGCGGTCCAATATCTTGATCGCCAGCTGGCTCTGGGAGACCTGCT 683
      ||||| |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
73 ArgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeule 89
864 GCACATCGCATTTGCATCCCTCATCAATGCTCTACAAGCTGCTGCAGAGG 733
      | ::::|::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

684 GCACATCGTCATTGCACATCCCTATCAATGCTACAGCTGCTGGCAGAGG 733

```
89 uLeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspA 106
734 ACTGCCCATTTGGAGCTGAGATGTGAAGTGGTGCCTTTTCATACAGAA 783
106 rgrTrpLeuPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLeu 122
784 CCTCCGTTGGGAACACACTGCTGAGTCTATGCTGCTCTGAGTATTGACAG 833
123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAspAr 139
834 ATATCGAGCTGTGTCTTCTTGAGTAGAATTAAAGGAATTGGGGTTCCAA 883
139 gTyrLysAlaIleValAlaGPrMetAspIleGlnAlaSerHisAlaLeu 156
884 AATGCACAGCAGTAGAATGTTTGGATTGGTGGTCTCTGCTGTTCTG 933
156 eTylsileCysLeuLysAlaAlaLeuIleTrpIleValSerMetLeuLeu 172
934 CTGTGCTCCTGAAGCCATAGTTTTCATATATATACGATGGACTACAAAGG 983
173 AlaIleProGluAlaValPheSerAspLeuHisProPheHisValLysAs 189
984 AAGTTATCTGCGAATC.....TGCTGCTTCATCCCGTTCAGAAAGACAG 1027
189 pThrAsnGlnThrPheIleSerCysAlaProTyrProHisSerAsnGluL 206
1028 CTTTCATGCAGTTTACAGACAGCAAAAGATTGGTGGCTGTTCAAGTTTC 1077
206 euHisProLysIleHisSerMetAlaSer.....PheLeuVal 218
1078 TATTCTCTGCTGCCATCGCCATCTGCTATTTTATACACTAATGAC 1127
219 PheTyrValIleProLeuAlaIleSerValTyrTyrPheIleAl 235
1128 CTGTCAAAATGTTGAGAAAGAAAGTGGCATGCAGATTGCTTTAAATGATC 1177
235 aArgAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIleH 252
1178 ACCTAAACAG.....AGACGGGAAGTGGCCCAAAACCGCTCTTT 1215
252 IsValLysLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeu 268
1216 TGCCGTGGCTGTGCTTTGCCCTCGCTGCTTCCCTTCCCTCAGCTCAGCAG 1265
269 ValPheValGlyLeuPheAlaPheCysThrLeuProAsnHisValIleTy 285
1266 GATTCTGAGCTCACTCTTTAATCAGAATGATCCCAATAGATGTGAAC 1315
285 rLeuTyrArgSerTyrHisTyrSerGluValAspThrSer.....M 299
1316 TTTTCAGCTTCTGTGTGTTGATGGACTATATTGGTATCAACATGCTTCA 1365
299 etLeuHisPhe.....ValThrSerIleCysAlaHisLeuLeuAlaPhe 313
1366 CTGAATTCCTCATTAACCACTGCTGCTGATTGTTGGTACCAAAAGATT 1415
314 ThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSerPh 330
1416 CAAAAACCTGCTTAAAGTCA...TGCTTATGCTGCTGGTCCAGTCAATTG 1462
330 eArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGlyLeuMeta 347
1463 AAGAAAAACAGTCTTGGAGGAAAGCAGTGTGCTTAAAGTTCAAAGCT 1512
347 snArgSerHisSerThrGlyArgSerThrThrCysMet..... 359
1513 AATGATCAGGNATATGACAACCTTCGTTCCAGTAAT 1548
360 .....ThrSerPheLysSerThrAsn 366
seq_name: /SIDS1/gcgcdata/geneseq/geneseq-emb1/AA2001.DAT: AAB66620
seq_documentation_block:
AAB66620 standard; protein; 384 AA.
AAB66620:
05-APR-2001 (first entry)
Mouse gastrin releasing peptide receptor.
Cholecystokinin; CCK receptor; purify.
Mus sp.
US6159173-B1.
02-JAN-2001.
10-MAR-1993; 9305-0029170.
07-FEB-1992; 92US-0831248.
01-APR-1992; 92US-0861769.
11-AUG-1992; 92US-0928033.
02-SEP-1992; 92US-0937603.
(USSH ) US DEPT HEALTH & HUMAN SERVICES.
Wank SA;
WPI; 2001-136725/14.
New cholecystokinin (CCK) receptor-encoding DNA molecule, useful for
producing and purifying human CCK receptor protein to
sequenceable-grade homogeneity -
Disclosure; Fig 3; 82pp; English.
The present invention relates to a cholecystokinin (CCK) receptor
protein. The CCK receptor-encoding DNA molecule is useful for
expressing and purifying CCK receptor protein to
sequenceable-grade homogeneity. The CCK receptor proteins
or fragments are useful for obtaining antibodies that can
recognize CCK-expressing cells. The transformed eukaryotic cell
lines are useful for studying the receptor in an environment
similar to its native environment, e.g. in the context of
studying the electrophysiology or binding properties of the receptor.
The transformed prokaryotic or insect cell line is useful for
expressing CCK receptor to produce large amounts of the receptor for
immunological purposes or for studying protein structure, e.g.
crystallography.
Sequence 384 AA:
alignment_scores:
Quality: 498.50 Length: 362
Ratio: 2.149 Gaps: 8
Percent Similarity: 64.088 Percent Identity: 33.702
alignment_block:
US-09-931-157-2 x AAB66620 ..
Align seg 1/1 to: AAB66620 from: 1 to: 384
484 CGCACATCTCCCTCCCTCCGTCGCAAGGACCCATCGAGATCAAGAGAC 533
25 GlnSerLeuSerProLysMetAspAsn.....TrpPheHisProGln 39
534 TTTCAATACATCAACACGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 583
39 yPheIleTyrValIleProAlaValTyrGlyLeuIleValIleGlyL 56
584 TCATCGGAATCCACACTTCTGAGAATTATCTACAAGAACAAAGTGCATG 633
56 euIleGlyAsnIleThrLeuIleLysIlePheCysThrValLysSerMet 72
```

534 CGAAACGGTCCCAATATCTTCATGCCAGCTTGGCTCTGGAGACCTGCT 683
 73 ArgAsnValProAsnLeuPheIleSerSerLeuAlaLeuGlyAspLeu 89
 684 GCACATCGCATTTGACATCCCTATCAATGTCTACAAGCTGTGGCAGAGG 733
 89 uLeuLeuValThrCysAlaProValAspAlaSerLysTyrLeuAlaAspA 106
 734 ACTGGCCATTGGAGCTGAGATGTGTAAGCTGGTGCCTTTCATACAGAAA 783
 106 rGTrPLeuPheGlyArgIleGlyCysLysLeuIleProPheIleGlnLeu 122
 784 GCCTCCGCGGAATCAGTGTGTGCTGATGTATGTCTGTGAGTATTGACAG 833
 123 ThrSerValGlyValSerValPheThrLeuThrAlaLeuSerAlaAspAr 139
 834 ATATCGAGCTGTCTCTCTGGAGTAGAATTAAGGAATTTGGGTTCCAA 883
 139 gTyrLysAlaIleValArgProMetAspIleGlnAlaSerHisAlaLeuM 156
 884 AATGGACAGCAGTAGAAATTTGTTGTTGTTGGTGTCTCTGTGGTCTG 933
 156 etLysIleCysLeuLysAlaAlaLeuIleTrpIleValSerMetLeuLeu 172
 934 GCTGTCTCTGAAGCCATAGTTTGTATATAATTACGATGGACTACAAAG 983
 173 AlaIleProGluAlaValPheSerAspLeuHisPropheHisValLysAs 189
 984 AAGTTATCTGCCAATCTGCTTCTTCATCCCTTCAGAGACAG 1027
 189 pThrAsnGlnThrPheIleSerCysAlaProTyrProHisSerAsnGluL 206
 1028 CTTTCATCGAGTTTTACAAGACAGCAAAAGATTGGTGGCTGTTCAGTTT 1077
 206 euHisProLysIleHisSerMetAlaSerPheLeuVal 218
 1078 TATTCTGCTGTCATGGCCATCGCATCTGCATTTTATATACACTAATGAC 1127
 219 PheTyrValIleProLeuAlaIleIleSerValTyrTyrPheIleAl 235
 1128 CTCTGAATGTTGAGAAAGAAAGTGGCATGTCAGATTGCTTTAAATGATC 1177
 235 aArgAsnLeuIleGlnSerAlaTyrAsnLeuProValGluGlyAsnIleH 252
 1178 ACCTAAACGAGAGACGGGAAGTGGCCAAACCGCTTT 1215
 252 IsValLysGlnIleGluSerArgLysArgLeuAlaLysThrValLeu 268
 1216 TGCTGTGCTCTGTCTTGGCTCTGTGCTTCCCTTCACCTCAGCAG 1265
 269 ValPheValGlyLeuPheAlaPheCysTrpLeuProAsnHisValIleTy 285
 1266 GATTCCTCAAGCTCACTCTTATATACAGATGATCCCAATAGATGTGAAC 1315
 285 rLeuTyrArgSerTyrHisTyrSerGluValAlaSpThrSerM 299
 1316 TTTTGAGCTTTCTGGTATGTTGACTATATTGGTATCAACATGGCTTCA 1365
 299 etLeuHisPheValThrSerIleCysAlaHisLeuLeuAlaPhe 313
 1366 CTGAATTCCTGATTAAACCAATTGCTCTGTATTTGGTGAGCAAAAGATT 1415
 314 ThrAsnSerCysValAsnProPheAlaLeuTyrLeuLeuSerLysSerPh 330
 1416 CAAAAACCTCTTAAAGTCA...TGCTATGCTGTGGTGCCAGTCATTG 1462
 330 eArgLysGlnPheAsnThrGlnLeuLeuCysCysGlnProGlyLeuMeta 347
 1463 AAGAAAAACAGCTCTGGAGGAAAAGCAGTCGTGCTTAAAGTTCAAAGCT 1512
 347 snArgSerHisSerThrGlyArgSerThrThrCysMet 359